GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:20:09 ; Search time 92.6283 Seconds
(without alignments)
1101.172 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1833.8	99.1	373	3	AAY53624	Aay53624 A bone ma
2	1833.8	99.1	373	3	AAY56969	Aay56969 Human MAG
3	1833.8	99.1	373	3	AAB24242	Aab24242 Human Nog
4	1833.8	99.1	373	4	AAB82350	Aab82350 Human NOG
5	1833.8	99.1	373	5	AAM47954	Aam47954 Human RTN
6	1833.8	99.1	373	5	ABP68601	Abp68601 Human pan
7	1833.8	99.1	373	5	ABB81079	Abb81079 Human neu
8	1825.8	98.7	373	5	ABG30937	Abg30937 Human Nog
9	1756.9	95.0	1192	3	AAY56967	Aay56967 Human MAG

10	1756.9	95.0	1192	4	AAB82349	Aab82349 Human NOG
11	1756.9	95.0	1192	4	AAU04591	Aau04591 Human Nog
12	1756.9	95.0	1192	5	ABG30938	Abg30938 Human Nog
13	1756.9	95.0	1192	5	ABP68600	Abp68600 Human pan
14	1756.9	95.0	1192	5	ABB81078	Abb81078 Human neu
15	1756.9	95.0	1192	6	ABR59667	Abr59667 Human Nog
16	1749.3	94.6	1178	3	AAY71311	Aay71311 Human neu
17	1556.2	84.1	379	7	ADB85283	Adb85283 Rat fooce
18	1551.6	83.9	361	3	AAY71385	Aay71385 Alternati
19	1549.3	83.7	359	3	AAY71558	Aay71558 Rat Nogo
20	1546.3	83.6	360	3	AAY71383	Aay71383 Rat neuri
21	1546.3	83.6	360	5	ABB81076	Abb81076 Rat neuro
22	1480.5	80.0	360	4	AAE03987	Aae03987 Human gen
23	1478.8	79.9	1163	3	AAY71310	Aay71310 Rat neuri
24	1478.8	79.9	1163	3	AAY71384	Aay71384 Alternati
25	1478.8	79.9	1163	5	ABB81074	Abb81074 Rat neuro
26	1476.8	79.8	1162	3	AAY71557	Aay71557 Rat Nogo
27	1475.2	79.7	1246	4	AAU33228	Aau33228 Novel hum
28	1324.8	71.6	291	4	AAM93484	Aam93484 Human pol
29	1028	55.6	893	3	AAY95012	Aay95012 Human sec
30	1028	55.6	983	6	ABU11573	Abu11573 Human MDD
31	993.5	53.7	642	2	AAW58383	Aaw58383 Human sec
32	993.5	53.7	642	4	AAB90682	Aab90682 Human BG1
33	967.3	52.3	403	3	AAY71563	Aay71563 Rat Nogo
34	938.2	50.7	200	4	AAB64514	Aab64514 Human sec
35	932.9	50.4	284	3	AAY95030	Aay95030 Human clo
36	927	50.1	199	2	AAW53947	Aaw53947 Human NSP
37	927	50.1	199	2	AAW78313	Aaw78313 Fragment
38	927	50.1	199	3	AAB12805	Aab12805 Human NSP
39	927	50.1	199	4	AAB82348	Aab82348 Human NOG
40	927	50.1	199	5	ABB81080	Abb81080 Human neu
41	918	49.6	199	2	AAY35903	Aay35903 Extended
42	906	49.0	199	3	AAY71559	Aay71559 Rat Nogo
43	904	48.9	199	5	ABB81077	Abb81077 Rat neuro
44	904	48.9	522	3	AAY71312	Aay71312 Rat neuri
45	886	47.9	199	5	ABG30939	Abg30939 Human Nog

ALIGNMENTS

RESULT 1
AAY53624
ID AAY53624 standard; protein; 373 AA.
XX
AC AAY53624;
XX
DT 22-FEB-2000 (first entry)
XX
DE A bone marrow secreted protein designated BMS112.
XX
KW Bone marrow secreted protein; bone marrow stromal cell; cytokine;
KW cell proliferation; cell differentiation; hematopoiesis; anaemia;
KW myeloid cell deficiency; lymphoid cell deficiency; myeloid cell;
KW erythroid progenitor cell; colony stimulating factor; granulocyte;
KW monocyte; macrophage; myelo-suppression; megakaryocyte; platelet;
KW platelet disorder; thrombocytopenia; hematopoeitic stem cell;

KW stem cell disorder; aplastic anaemia; bone differentiation;
KW paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon;
KW ligament; nerve; wound healing; tissue repair; burn; incision; ulcer;
KW bone fracture; cartilage damage; artificial joint.
XX
OS Homo sapiens.
XX
PN WO9933979-A2.
XX
PD 08-JUL-1999.
XX
PF 18-DEC-1998; 98WO-US027008.
XX
PR 30-DEC-1997; 97US-0068958P.
PR 24-SEP-1998; 98US-0101603P.
PR 30-SEP-1998; 98US-0102540P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Lin H, Cao L;
XX
DR WPI; 2000-038344/03.
DR N-PSDB; AAZ36230.
XX
PT New isolated human polynucleotide and secreted proteins can induce
PT production of other cytokines in certain cell populations.
XX
PS Claim 2; Page 74; 120pp; English.
XX
CC AAY53622-43 represent bone marrow secreted proteins of human bone marrow
CC stromal cells. The proteins can exhibit cytokine, cell proliferation, or
CC cell differentiation activity (either inducing or inhibiting). They can
CC be used to support colony forming cells or factor-dependent cell lines,
CC to regulate hematopoiesis, and to treat myeloid or lymphoid cell
CC deficiencies. In addition, they may be used to support the growth and
CC proliferation of erythroid progenitor cells, and to treat various
CC anaemias. They can have colony stimulating factor (CSF) activity and can
CC be used to support the growth and proliferation of myeloid cells such as
CC granulocytes, monocytes or macrophages, to prevent or treat myelo-
CC suppression, to support the growth and proliferation of megakaryocytes
CC and platelets, thereby allowing prevention or treatment of platelet
CC disorders such as thrombocytopenia, to support the growth and
CC proliferation of hematopoeitic stem cells, either in place of or in
CC conjunction with platelet transfusions, to treat stem cell disorders,
CC such as aplastic anaemia and paroxysmal nocturnal hemoglobinuria, or to
CC repopulate the stem cell compartment after irradiation or chemotherapy.
CC They can be used for growth or differentiation of bone, cartilage,
CC tendon, ligament, or nerve tissue, as well as for wound healing and
CC tissue repair and replacement, and in the treatment of burns, incisions
CC and ulcers, to induce cartilage and/or bone growth in circumstances where
CC bone is not normally formed and thus have an application in healing bone
CC fractures and cartilage damage or defects, prophylactic use in fracture
CC reduction and also in the improved fixation of artificial joints
XX
SQ Sequence 373 AA;

Query Match 99.1%; Score 1833.8; DB 3; Length 373;

```
  Best Local Similarity   96.5%;  Pred. No. 1.5e-64;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

```
RESULT 2
AAY56969
ID   AAY56969 standard; protein; 373 AA.
XX
AC   AAY56969;
XX
DT   25-APR-2000  (first entry)
XX
DE   Human MAGI polypeptide variant.
XX
KW   MAGI protein; neuroendocrine-specific protein; neuropathy; human;
KW   spinal injury; neuronal degeneration; neuromuscular disorder; cancer;
KW   psychiatric disorder; developmental disorder; inflammatory disorder;
KW   stroke; cytostatic; cerebroprotective; neuroprotective; variant.
XX
OS   Homo sapiens.
XX
PN   WO200005364-A1.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-GB002360.
XX
PR   22-JUL-1998;   98GB-00016024.
PR   19-JUL-1999;   99GB-00016898.
```

XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Prinjha RK;
XX
DR WPI; 2000-182693/16.
DR N-PSDB; AAZ56888.
XX
PT Novel polypeptides related to neuroendocrine-specific proteins and
PT polynucleotides useful for diagnosis of various diseases and for
PT treatment of cancer and neurological disorders.
XX
PS Claim 2; Page 22; 35pp; English.
XX
CC The invention relates to human MAGI protein, which is similar to
CC neuroendocrine-specific protein. The MAGI protein can be expressed by
CC standard recombinant methodology. The MAGI polypeptides, polynucleotides
CC and antibodies are useful for treating diseases, including neuropathies,
CC spinal injury, neuronal degeneration, neuromuscular disorders,
CC psychiatric disorders and developmental disorders, cancer, stroke and
CC inflammatory disorders. The polynucleoitde is also useful for chromosome
CC localization and for tissue expression studies. The present sequence
CC represents the human MAGI protein variant
XX
SQ Sequence 373 AA;

Query Match 99.1%; Score 1833.8; DB 3; Length 373;
Best Local Similarity 96.5%; Pred. No. 1.5e-64;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 3
AAB24242
ID AAB24242 standard; protein; 373 AA.
XX
AC AAB24242;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human Nogo B protein sequence SEQ ID NO:2.
XX
KW Human; Nogo B; cell stress response; hyperphosphorylated; brain tumour;
KW stress-phosphorylated endoplasmic reticulum protein; cytostatic;
KW gene therapy; cell growth; cellular stress response; neuron growth;
KW regulator of oxidative stress; inhibitor of neurite outgrowth;
KW axon regeneration; diagnosis; cancer; identification.
XX
OS Homo sapiens.
XX
PN WO200060083-A1.
XX
PD 12-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US009383.
XX
PR 08-APR-1999; 99US-0128372P.
PR 21-JUN-1999; 99US-0140331P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Wei D, Halenbeck R, Williams LT;
XX
DR WPI; 2000-665007/64.
DR N-PSDB; AAC64406.
XX
PT Novel protein associated with cell stress response useful for modulating
PT stress levels, cell growth, diagnosis and treatment of cancer and
PT malignant growth and for identifying agonists and antagonists.
XX
PS Claim 11; Fig 1A; 68pp; English.
XX
CC The present sequence represents a human stress-phosphorylated endoplasmic
CC reticulum protein, designated Nogo B. Nogo B has cytostatic activity and
CC is a modulator of the storage and exchange of calcium, cell growth and
CC cellular stress response. It can: regulate oxidative stress; inhibit
CC neurite outgrowth, neuron growth and axon regeneration. Nogo B
CC polypeptides and polynucleotides are useful for modulating stress levels
CC and cellular stress-response, cell growth and viability, diagnosis and
CC treatment of cancer, malignant growth and other Nogo B related diseases.
CC Nogo B polypeptides are also useful to screen combinatorial libraries to
CC identify agonist or antagonist. Antibodies against Nogo B polypeptides
CC are useful for affinity chromatography and distinguishing Nogo B
CC polypeptides
XX
SQ Sequence 373 AA;

```
  Query Match             99.1%;  Score 1833.8;  DB 3;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 1.5e-64;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373


RESULT 4
AAB82350
ID   AAB82350 standard; protein; 373 AA.
XX
AC   AAB82350;
XX
DT   23-JUL-2001  (first entry)
XX
DE   Human NOGO-B protein.
XX
KW   NOGO-B; human; chromosome 2p21; neuropathy; spinal injury; brain injury;
KW   stroke; neuronal degeneration; Alzheimer's disease; Parkinson's disease;
KW   neuromuscular disorder; psychiatric disorder; developmental disorder;
KW   neuroprotective; nootropic; neuroleptic; antiparkinsonian;
KW   cerebroprotective; neuroleptic; diagnosis; therapy.
XX
OS   Homo sapiens.
XX
PN   WO200136631-A1.
XX
PD   25-MAY-2001.
XX
PF   14-NOV-2000; 2000WO-GB004345.
XX
```

PR 15-NOV-1999; 99GB-00026995.
PR 24-JAN-2000; 2000GB-00001550.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Prinjha R;
XX
DR WPI; 2001-343822/36.
DR N-PSDB; AAF90325.
XX
PT New polypeptide designated NOGO-C is a splice variant of the human NOGO
PT gene and may be useful in the treatment of neural disorders including
PT Alzheimer's and Parkinson's diseases.
XX
PS Disclosure; Page 27-38; 25pp; English.
XX
CC The present sequence is that of human NOGO-B. NOGO-B is a previously
CC known splice variant of the human NOGO gene on chromosome 2p21. The
CC invention relates to a novel splice variant, NOGO-C (see AAB82348). It
CC provides NOGO-C polypeptides and polynucleotides, and methods for
CC producing such polypeptides by recombinant techniques. Also disclosed are
CC methods for utilising NOGO-C polypeptides and polynucleotides in the
CC treatment of diseases including neuropathies, spinal injury, brain
CC injury, stroke, neuronal degeneration, for example Alzheimer's disease
CC and Parkinson's disease, neuromuscular disorders, psychiatric disorders
CC and developmental disorders. Also provided are methods for identifying
CC agonists and agonists for use in treating conditions associated with NOGO
CC -C imbalance, and diagnostic assays for detecting diseases associated
CC with inappropriate NOGO-C activity or levels
XX
SQ Sequence 373 AA;

Query Match 99.1%; Score 1833.8; DB 4; Length 373;
Best Local Similarity 96.5%; Pred. No. 1.5e-64;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 5
AAM47954
ID AAM47954 standard; protein; 373 AA.
XX
AC AAM47954;
XX
DT 04-MAR-2002 (first entry)
XX
DE Human RTN4B SEQ ID NO 4.
XX
KW Human; RTN4B.
XX
OS Homo sapiens.
XX
PN CN1311439-A.
XX
PD 05-SEP-2001.
XX
PF 02-MAR-2000; 2000CN-00111791.
XX
PR 02-MAR-2000; 2000CN-00111791.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Fu Q, Zhao Y;
XX
DR WPI; 2002-049934/07.
DR N-PSDB; ABA05903.
XX
PT Human RTN 4B protein and coding sequence, its preparation and use.
XX
PS Claim 5; Page 11 (Disclosure); 27pp; Chinese.
XX
CC The invention relates to human RTN4B protein and coding sequence, useful
CC for providing a cDNA sequence of human RTN4B. The protein is an isomer of
CC RTN4 with RTN family members. The present invention also refers to
CC polypeptide encoded by a nucleotide sequence and manufacturing method and
CC application of the polypeptide and polynucleotide
XX
SQ Sequence 373 AA;

```
  Query Match             99.1%;  Score 1833.8;  DB 5;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 1.5e-64;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
```

```
Qy          61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy         121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy         174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy         229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy         289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy         349 IQAKIPGLKRKAE 361
               |||||||||||||
Db         361 IQAKIPGLKRKAE 373
```

RESULT 6
ABP68601
ID   ABP68601 standard; protein; 373 AA.
XX
AC   ABP68601;
XX
DT   14-JAN-2003 (first entry)
XX
DE   Human pancreatic cancer expressed protein SEQ ID NO 72.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US002781.
XX
PR   30-JAN-2001; 2001US-0265305P.
PR   31-JAN-2001; 2001US-0265682P.
PR   09-FEB-2001; 2001US-0267568P.
PR   21-MAR-2001; 2001US-0278651P.
PR   28-APR-2001; 2001US-0287112P.
PR   16-MAY-2001; 2001US-0291631P.
PR   12-JUL-2001; 2001US-0305484P.
PR   20-AUG-2001; 2001US-0313999P.
PR   27-NOV-2001; 2001US-0333626P.
XX
PA   (CORI-) CORIXA CORP.

XX
PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR WPI; 2002-627435/67.
DR N-PSDB; ABV94681.
XX
PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT diagnosing, preventing and/or treating cancer, particularly pancreatic
PT cancer.
XX
PS Claim 2; SEQ ID NO 72; 300pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC complements of (a); (c) sequences consisting of at least 20 contiguous
CC residues of (a); (d) sequences that hybridize to (a), under moderately
CC stringent conditions; (e) sequences having at least 75% or 90% identity
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC in a patient and compositions comprising polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells expressing the polypeptide are useful in treating pancreatic cancer
CC and stimulating an immune response. The polynucleotides can be used as
CC probes or primers for nucleic acid hybridisation, in the design and
CC preparation of ribozyme molecules for inhibiting expression of the tumour
CC polypeptides and proteins in the tumour cells, in vaccines and for gene
CC therapy. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 373 AA;

Query Match 99.1%; Score 1833.8; DB 5; Length 373;
Best Local Similarity 96.5%; Pred. No. 1.5e-64;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
                    |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 7
ABB81079
ID ABB81079 standard; protein; 373 AA.
XX
AC ABB81079;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human neurotransmitter receptor protein Nogo-B.
XX
KW Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;
KW central nervous system; peripheral nervous system; tranquillizer; Nogo;
KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;
KW nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;
KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;
KW neurotransmitter receptor; human; receptor.
XX
OS Homo sapiens.
XX
PN US2002072493-A1.
XX
PD 13-JUN-2002.
XX
PF 28-JUN-2001; 2001US-00893348.
XX
PR 19-MAY-1998; 98IL-00124500.
PR 21-JUL-1998; 98WO-US014715.
PR 22-DEC-1998; 98US-00218277.
PR 19-MAY-1999; 99US-00314161.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonego A;
PI Moalem G;
XX
DR WPI; 2002-607255/65.
DR N-PSDB; ABN86601.
XX
PT Promoting nerve regeneration and preventing neuronal degeneration in the
PT central/peripheral nervous system from injury/disease, comprises
PT administering nervous system-specific activated T cells/antigen, or
PT analogs/peptides.
XX
PS Example; Page 56-57; 93pp; English.
XX
CC The invention relates to promoting nerve regeneration or conferring
CC neuroprotection and preventing or inhibiting neuronal degeneration in the
CC central/peripheral nervous system (NS). The method involves administering
CC NS-specific activated T cells, NS-specific antigen, its analogue or its

CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or
CC combinations. The method is useful for promoting nerve regeneration and
CC preventing neuronal degeneration in central/peripheral nervous system
CC from injury/disease, where the injury is spinal cord injury, blunt
CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or
CC damages caused by surgery such as tumour excision. The disease is not an
CC autoimmune disease or neoplasm. The disease results in a degenerative
CC process occurring in either gray or white matter or both. The disease is
CC diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's
CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,
CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and
CC vitamin deficiency, intervertebral disc herniation, prion diseases such
CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral
CC neuropathies associated with various diseases, including but not limited
CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute
CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary
CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption
CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-
CC pathies, complications of various drugs (e.g., metronidazole) and toxins
CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia
CC telangectasia, Friedreich's ataxia, amyloid polyneuropathies,
CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's
CC disease, or lipoproteinemia. The present sequence represents the human
CC neurotransmitter receptor protein Nogo-B, an example of NS-specific
CC antigen
XX
SQ Sequence 373 AA;

Query Match 99.1%; Score 1833.8; DB 5; Length 373;
Best Local Similarity 96.5%; Pred. No. 1.5e-64;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
```

```
                    ||||||||||||
Db          361 IQAKIPGLKRKAE 373


RESULT 8
ABG30937
ID   ABG30937 standard; protein; 373 AA.
XX
AC   ABG30937;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human NogoB protein.
XX
KW   Human; Nogo; BACE; acute neuronal injury; spinal injury; head injury;
KW   stroke; peripheral nerve damage; neoplastic disorder; glioblastoma;
KW   neuroblastoma; hyperproliferative disorder; dysproliferative disorder;
KW   cirrhosis; psoriasis; keloid formation; fibrocystic condition; cancer;
KW   tissue hypertrophy; central nervous system; axon regeneration; NogoB;
KW   Nogo-associated disease; metastasis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         92. .104
FT                   /note= "NogoB (ASY) peptide"
FT   Peptide         180. .193
FT                   /note= "NogoB (ASY) peptide, distinguishes NogoB from
FT                   other Nogo isoforms"
FT   Peptide         272. .285
FT                   /note= "NogoB (ASY) peptide"
XX
PN   WO200257483-A2.
XX
PD   25-JUL-2002.
XX
PF   18-JAN-2002; 2002WO-GB000228.
XX
PR   18-JAN-2001; 2001GB-00001312.
XX
PA   (GLAX ) GLAXO GROUP LTD.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Blackstock WP,  Hale RS,  Prinjha R,  Rowley A;
XX
DR   WPI; 2002-599722/64.
DR   N-PSDB; ABK90133.
XX
PT   Identifying modulators of Nogo or BACE activity for treating acute
PT   neuronal injuries, neoplastic or dysproliferative disorders, comprises
PT   providing and monitoring interaction between Nogo and BACE polypeptides.
XX
PS   Example 1; Page 25; 68pp; English.
XX
CC   The present invention relates to a new method of identifying modulators
CC   of Nogo function or BACE activity. The method involves providing Nogo and
CC   BACE polypeptides capable of binding with each other, monitoring the
```

CC interaction between these polypeptides, and determining if the test agent
CC is a modulator of Nogo or BACE activity. The method is useful in treating
CC acute neuronal injuries, such as spinal or head injury, stroke,
CC peripheral nerve damage, and in neoplastic (e.g. glioblastomas,
CC neuroblastomas), hyperproliferative or dysproliferative disorders (e.g.
CC cirrhosis, psoriasis, keloid formation, fibrocystic conditions, tissue
CC hypertrophy) of the central nervous system. The BACE polypeptide is
CC useful in screening methods to identify agents that may act as modulators
CC of BACE activity and in particular agents that may be useful in treating
CC Nogo-associated diseases. The modulators of Nogo or BACE polypeptides,
CC and the polynucleotide encoding the BACE polypeptide are useful in
CC manufacturing a medicament for the treatment or prevention of disorders
CC responsive to the modulation of Nogo activity, in alleviating the
CC symptoms or improving the condition of a patient suffering from this
CC disorder, in axon regeneration, or in preventing metastasis or spreading
CC of a cancer. The polynucleotide may also be an essential component in
CC assays, a probe, in recombinant protein synthesis, and in gene therapy
CC techniques. The present amino acid sequence represents the human NogoB
CC protein of the invention
XX
SQ Sequence 373 AA;

Query Match 98.7%; Score 1825.8; DB 5; Length 373;
Best Local Similarity 96.0%; Pred. No. 3e-64;
Matches 358; Conservative 1; Mismatches 2; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPCWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              ||||||||||||||:||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPQDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 9
AAY56967

ID AAY56967 standard; protein; 1192 AA.
XX
AC AAY56967;
XX
DT 25-APR-2000 (first entry)
XX
DE Human MAGI polypeptide.
XX
KW MAGI protein; neuroendocrine-specific protein; neuropathy; human;
KW spinal injury; neuronal degeneration; neuromuscular disorder; cancer;
KW psychiatric disorder; developmental disorder; inflammatory disorder;
KW stroke; cytostatic; cerebroprotective; neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200005364-A1.
XX
PD 03-FEB-2000.
XX
PF 21-JUL-1999; 99WO-GB002360.
XX
PR 22-JUL-1998; 98GB-00016024.
PR 19-JUL-1999; 99GB-00016898.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Prinjha RK;
XX
DR WPI; 2000-182693/16.
DR N-PSDB; AAZ56886.
XX
PT Novel polypeptides related to neuroendocrine-specific proteins and
PT polynucleotides useful for diagnosis of various diseases and for
PT treatment of cancer and neurological disorders.
XX
PS Claim 2; Page 20-21; 35pp; English.
XX
CC The invention relates to human MAGI protein, which is similar to
CC neuroendocrine-specific protein. The MAGI protein can be expressed by
CC standard recombinant methodology. The MAGI polypeptides, polynucleotides
CC and antibodies are useful for treating diseases, including neuropathies,
CC spinal injury, neuronal degeneration, neuromuscular disorders,
CC psychiatric disorders and developmental disorders, cancer, stroke and
CC inflammatory disorders. The polynucleoitde is also useful for chromosome
CC localization and for tissue expression studies. The present sequence
CC represents the human MAGI protein
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 3; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

Qy 1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

```
Qy        61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy       121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy       172 ------------------------------------------------------------ 171

Db       181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy       172 ------------------------------------------------------------ 171

Db       241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy       172 ------------------------------------------------------------ 171

Db       301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy       172 ------------------------------------------------------------ 171

Db       361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy       172 ------------------------------------------------------------ 171

Db       421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy       172 ------------------------------------------------------------ 171

Db       481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy       172 ------------------------------------------------------------ 171

Db       541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy       172 ------------------------------------------------------------ 171

Db       601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy       172 ------------------------------------------------------------ 171

Db       661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy       172 ------------------------------------------------------------ 171

Db       721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy       172 ------------------------------------------------------------ 171

Db       781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy       172 ------------------------------------------------------------ 171

Db       841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy       172 ------------------------------------------------------------ 171
```

```
Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                     ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 10
AAB82349
ID AAB82349 standard; protein; 1192 AA.
XX
AC AAB82349;
XX
DT 23-JUL-2001 (first entry)
XX
DE Human NOGO-A protein.
XX
KW NOGO-A; human; chromosome 2p21; neuropathy; spinal injury; brain injury;
KW stroke; neuronal degeneration; Alzheimer's disease; Parkinson's disease;
KW neuromuscular disorder; psychiatric disorder; developmental disorder;
KW neuroprotective; nootropic; neuroleptic; antiparkinsonian;
KW cerebroprotective; neuroleptic; diagnosis; therapy.
XX
OS Homo sapiens.
XX
PN WO200136631-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-GB004345.
XX
PR 15-NOV-1999; 99GB-00026995.
PR 24-JAN-2000; 2000GB-00001550.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Prinjha R;
XX
DR WPI; 2001-343822/36.
DR N-PSDB; AAF90324.
XX
PT New polypeptide designated NOGO-C is a splice variant of the human NOGO
PT gene and may be useful in the treatment of neural disorders including
PT Alzheimer's and Parkinson's diseases.

XX
PS Disclosure; Page 26-27; 25pp; English.
XX
CC The present sequence is that of human NOGO-A. NOGO-A is a previously
CC known splice variant of the human NOGO gene on chromosome 2p21. The
CC invention relates to a novel splice variant, NOGO-C (see AAB82348). It
CC provides NOGO-C polypeptides and polynucleotides, and methods for
CC producing such polypeptides by recombinant techniques. Also disclosed are
CC methods for utilising NOGO-C polypeptides and polynucleotides in the
CC treatment of diseases including neuropathies, spinal injury, brain
CC injury, stroke, neuronal degeneration, for example Alzheimer's disease
CC and Parkinson's disease, neuromuscular disorders, psychiatric disorders
CC and developmental disorders. Also provided are methods for identifying
CC agonists and agonists for use in treating conditions associated with NOGO
CC -C imbalance, and diagnostic assays for detecting diseases associated
CC with inappropriate NOGO-C activity or levels
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 4; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171
```

```
Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

```
RESULT 11
AAU04591
ID   AAU04591 standard; protein; 1192 AA.
XX
AC   AAU04591;
XX
DT   26-SEP-2001  (first entry)
XX
DE   Human Nogo protein.
XX
```

KW Human; Nogo receptor; axonal growth; immunogen; antibody; nogo protein;
KW cranial trauma; cerebral trauma; spinal cord injury; stroke;
KW demyelinating disease; multiple sclerosis; monophasis demyelination;
KW encephalomyelitis; multifocal leukoencephalopathy; panencephalitis;
KW Marchiafava-Bignami disease; pontine myelinolysis; adrenoleukodystrophy;
KW Pelizaeus-Merzbacher disease; Spongy degeneration; Alexander's disease;
KW Canavan's disease; metachromatic leukodystrophy; viral infection;
KW Krabbe's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1054. .1119
FT /label= Lumenal_extracellular_domain
FT /note= "This sequence is specifically claimed"
FT Peptide 1055. .1094
FT /label= Pep1
FT /note= "Receptor binding inhibitory peptide. This
FT sequence is specifically claimed"
FT Peptide 1064. .1088
FT /label= Pep2
FT /note= "Receptor binding inhibitory peptide. This
FT sequence is specifically claimed"
FT Peptide 1074. .1098
FT /label= Pep3
FT /note= "Receptor binding inhibitory peptide. This
FT sequence is specifically claimed"
FT Peptide 1084. .1108
FT /label= Pep4
FT /note= "Receptor binding inhibitory peptide. This
FT sequence is specifically claimed"
FT Peptide 1095. .1119
FT /label= Pep5
FT /note= "Receptor binding inhibitory peptide. This
FT sequence is specifically claimed"
XX
PN WO200151520-A2.
XX
PD 19-JUL-2001.
XX
PF 12-JAN-2001; 2001WO-US001041.
XX
PR 12-JAN-2000; 2000US-0175707P.
PR 26-MAY-2000; 2000US-0207366P.
PR 29-SEP-2000; 2000US-0236378P.
XX
PA (UYYA) UNIV YALE.
XX
PI Strittmatter SM;
XX
DR WPI; 2001-442138/47.
DR N-PSDB; AAS09453.
XX
PT Novel Nogo receptor protein useful for identifying modulator of Nogo
PT protein or Nogo receptor protein, which is useful for treating central
PT nervous system disorders.
XX

PS Example 1; Page 101-104; 109pp; English.
XX
CC The sequence is the human Nogo protein, a 250kDa myelin-associated axon
CC growth inhibitor. The invention relates to the use of the nogo receptor,
CC nogo protein, their nucleic acids, vectors expressing them and antibodies
CC against them, to isolate agents which block nogo receptor mediated axonal
CC growth. The agent is useful for treating a central nervous system
CC disorder which is a result of cranial or cerebral trauma, spinal cord
CC injury, stroke or a demyelinating disease selected from multiple
CC sclerosis, monophasis demyelination, encephalomyelitis, multifocal
CC leukoencephalopathy, panencephalitis, Marchiafava-Bignami disease,
CC pontine myelinolysis, adrenoleukodystrophy, Pelizaeus-Merzbacher disease,
CC Spongy degeneration, Alexander's disease, Canavan's disease,
CC metachromatic leukodystrophy, viral infection and Krabbe's disease
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 4; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

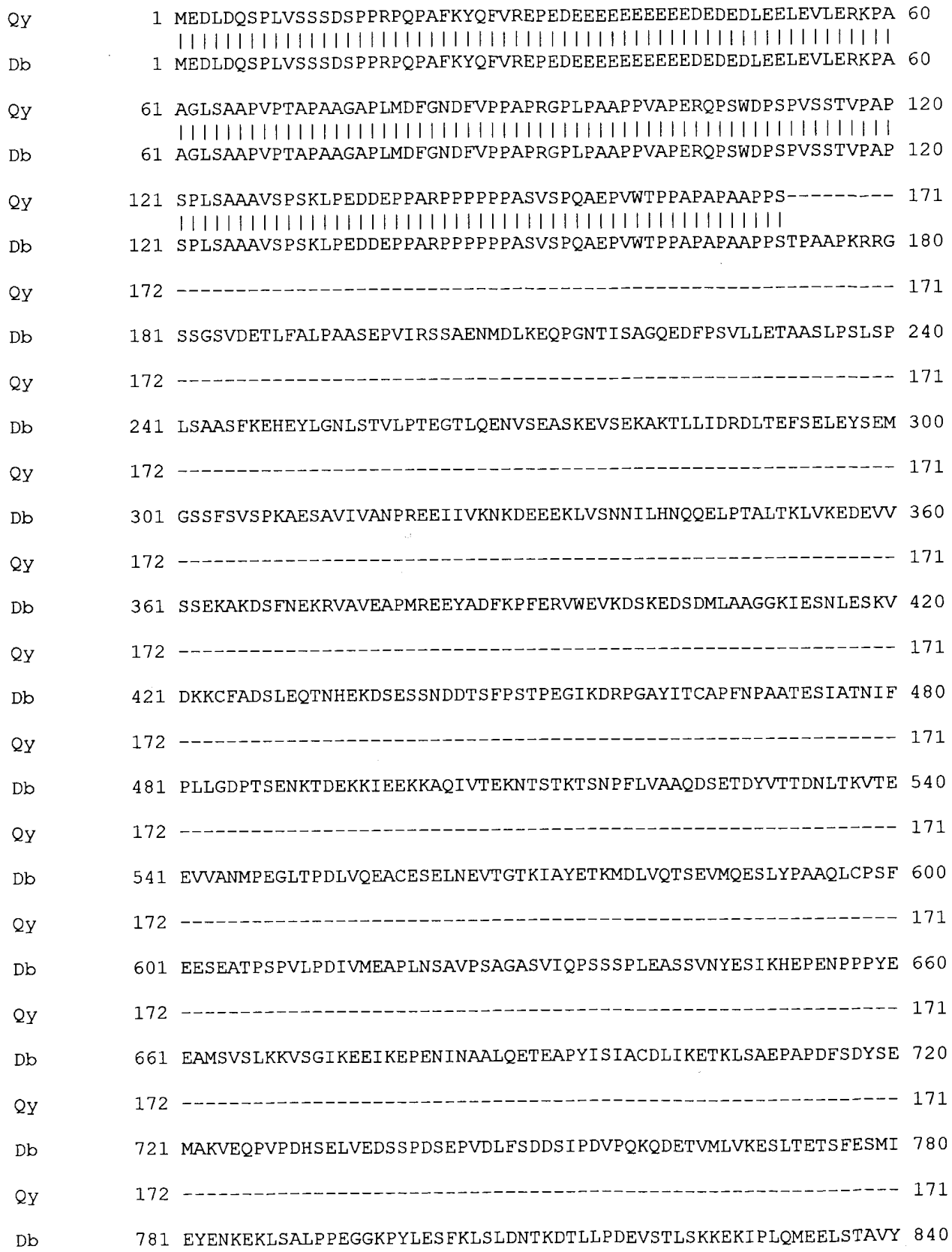
```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540
```

```
Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 12
ABG30938
ID ABG30938 standard; protein; 1192 AA.
XX
AC ABG30938;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NogoA protein.
XX
KW Human; Nogo; BACE; acute neuronal injury; spinal injury; head injury;
KW stroke; peripheral nerve damage; neoplastic disorder; glioblastoma;

KW neuroblastoma; hyperproliferative disorder; dysproliferative disorder;
KW cirrhosis; psoriasis; keloid formation; fibrocystic condition; cancer;
KW tissue hypertrophy; central nervous system; axon regeneration; NogoA;
KW Nogo-associated disease; metastasis.
XX
OS Homo sapiens.
XX
PN WO200257483-A2.
XX
PD 25-JUL-2002.
XX
PF 18-JAN-2002; 2002WO-GB000228.
XX
PR 18-JAN-2001; 2001GB-00001312.
XX
PA (GLAX) GLAXO GROUP LTD.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Blackstock WP, Hale RS, Prinjha R, Rowley A;
XX
DR WPI; 2002-599722/64.
DR N-PSDB; ABK90134.
XX
PT Identifying modulators of Nogo or BACE activity for treating acute
PT neuronal injuries, neoplastic or dysproliferative disorders, comprises
PT providing and monitoring interaction between Nogo and BACE polypeptides.
XX
PS Disclosure; Page 59-62; 68pp; English.
XX
CC The present invention relates to a new method of identifying modulators
CC of Nogo function or BACE activity. The method involves providing Nogo and
CC BACE polypeptides capable of binding with each other, monitoring the
CC interaction between these polypeptides, and determining if the test agent
CC is a modulator of Nogo or BACE activity. The method is useful in treating
CC acute neuronal injuries, such as spinal or head injury, stroke,
CC peripheral nerve damage, and in neoplastic (e.g. glioblastomas,
CC neuroblastomas), hyperproliferative or dysproliferative disorders (e.g.
CC cirrhosis, psoriasis, keloid formation, fibrocystic conditions, tissue
CC hypertrophy) of the central nervous system. The BACE polypeptide is
CC useful in screening methods to identify agents that may act as modulators
CC of BACE activity and in particular agents that may be useful in treating
CC Nogo-associated diseases. The modulators of Nogo or BACE polypeptides,
CC and the polynucleotide encoding the BACE polypeptide are useful in
CC manufacturing a medicament for the treatment or prevention of disorders
CC responsive to the modulation of Nogo activity, in alleviating the
CC symptoms or improving the condition of a patient suffering from this
CC disorder, in axon regeneration, or in preventing metastasis or spreading
CC of a cancer. The polynucleotide may also be an essential component in
CC assays, a probe, in recombinant protein synthesis, and in gene therapy
CC techniques. The present amino acid sequence represents the human NogoA
CC protein of the invention
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 5; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;


Qy 1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
Db 1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
Qy 61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
Db 61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
Qy 121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
Db 121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
Qy 172 --------- 171
Db 181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240
Qy 172 --------- 171
Db 241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300
Qy 172 --------- 171
Db 301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360
Qy 172 --------- 171
Db 361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420
Qy 172 --------- 171
Db 421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480
Qy 172 --------- 171
Db 481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540
Qy 172 --------- 171
Db 541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
Qy 172 --------- 171
Db 601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
Qy 172 --------- 171
Db 661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720
Qy 172 --------- 171
Db 721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780
Qy 172 --------- 171
Db 781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

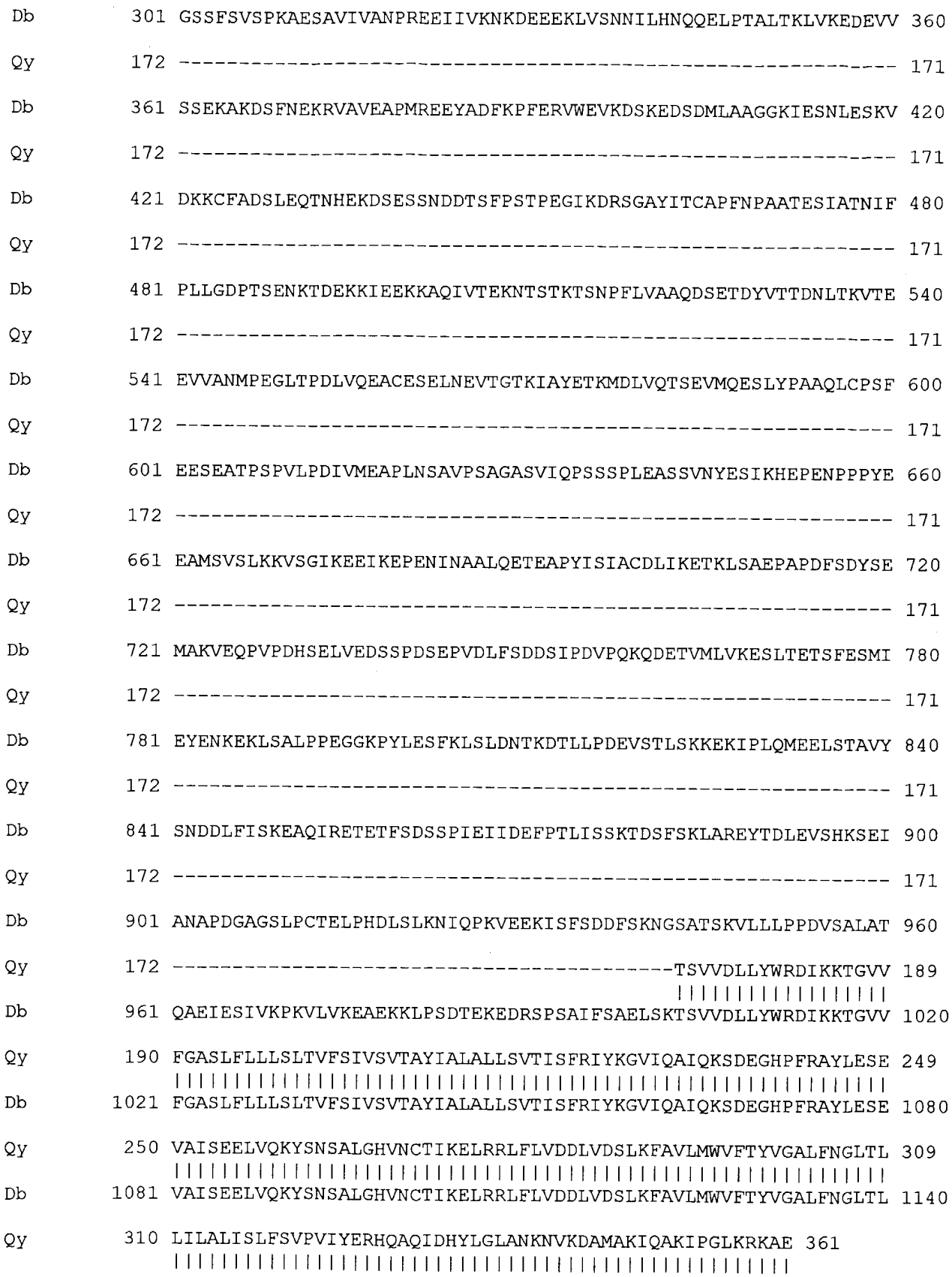
```
Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                      ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

```
RESULT 13
ABP68600
ID   ABP68600 standard; protein; 1192 AA.
XX
AC   ABP68600;
XX
DT   14-JAN-2003  (first entry)
XX
DE   Human pancreatic cancer expressed protein SEQ ID NO 71.
XX
KW   Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW   cytostatic; tumour.
XX
OS   Homo sapiens.
XX
PN   WO200260317-A2.
XX
PD   08-AUG-2002.
XX
PF   30-JAN-2002; 2002WO-US002781.
XX
PR   30-JAN-2001; 2001US-0265305P.
PR   31-JAN-2001; 2001US-0265682P.
PR   09-FEB-2001; 2001US-0267568P.
PR   21-MAR-2001; 2001US-0278651P.
PR   28-APR-2001; 2001US-0287112P.
PR   16-MAY-2001; 2001US-0291631P.
PR   12-JUL-2001; 2001US-0305484P.
PR   20-AUG-2001; 2001US-0313999P.
PR   27-NOV-2001; 2001US-0333626P.
XX
PA   (CORI-) CORIXA CORP.
```

XX
PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR WPI; 2002-627435/67.
DR N-PSDB; ABV94680.
XX
PT New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT diagnosing, preventing and/or treating cancer, particularly pancreatic
PT cancer.
XX
PS Claim 2; SEQ ID NO 71; 300pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising: (a)
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
CC complements of (a); (c) sequences consisting of at least 20 contiguous
CC residues of (a); (d) sequences that hybridize to (a), under moderately
CC stringent conditions; (e) sequences having at least 75% or 90% identity
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
CC in a patient and compositions comprising polypeptides, polynucleotides,
CC antibodies, fusion proteins, T cell populations and antigen presenting
CC cells expressing the polypeptide are useful in treating pancreatic cancer
CC and stimulating an immune response. The polynucleotides can be used as
CC probes or primers for nucleic acid hybridisation, in the design and
CC preparation of ribozyme molecules for inhibiting expression of the tumour
CC polypeptides and proteins in the tumour cells, in vaccines and for gene
CC therapy. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 5; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171
```


Db 301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360
Qy 172 ---------- 171
Db 361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420
Qy 172 ---------- 171
Db 421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480
Qy 172 ---------- 171
Db 481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540
Qy 172 ---------- 171
Db 541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
Qy 172 ---------- 171
Db 601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
Qy 172 ---------- 171
Db 661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720
Qy 172 ---------- 171
Db 721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780
Qy 172 ---------- 171
Db 781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840
Qy 172 ---------- 171
Db 841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900
Qy 172 ---------- 171
Db 901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960
Qy 172 ----------TSVVDLLYWRDIKKTGVV 189
Db 961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020
Qy 190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
Db 1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080
Qy 250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
Db 1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140
Qy 310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361

Db 1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192

RESULT 14
ABB81078
ID ABB81078 standard; protein; 1192 AA.
XX
AC ABB81078;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human neurotransmitter receptor protein Nogo-A.
XX
KW Nerve regeneration; neuroprotection; neuronal degeneration; CNS; PNS;
KW central nervous system; peripheral nervous system; tranquillizer; Nogo;
KW vulnerary; cerebroprotective; anti-tumour; antidiabetic; anticonvulsant;
KW nootropic; antiparkinsonian; ophthalmological; analgesic; hepatotropic;
KW osteopathic; vasotropic; nephrotropic; cytostatic; antigen; gene therapy;
KW neurotransmitter receptor; human; receptor.
XX
OS Homo sapiens.
XX
PN US2002072493-A1.
XX
PD 13-JUN-2002.
XX
PF 28-JUN-2001; 2001US-00893348.
XX
PR 19-MAY-1998; 98IL-00124500.
PR 21-JUL-1998; 98WO-US014715.
PR 22-DEC-1998; 98US-00218277.
PR 19-MAY-1999; 99US-00314161.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Eisenbach-Schwartz M, Hauben E, Cohen IR, Beserman P, Mosonego A;
PI Moalem G;
XX
DR WPI; 2002-607255/65.
DR N-PSDB; ABN86601.
XX
PT Promoting nerve regeneration and preventing neuronal degeneration in the
PT central/peripheral nervous system from injury/disease, comprises
PT administering nervous system-specific activated T cells/antigen, or
PT analogs/peptides.
XX
PS Example; Page 53-56; 93pp; English.
XX
CC The invention relates to promoting nerve regeneration or conferring
CC neuroprotection and preventing or inhibiting neuronal degeneration in the
CC central/peripheral nervous system (NS). The method involves administering
CC NS-specific activated T cells, NS-specific antigen, its analogue or its
CC peptide, a nucleotide sequence the NS-specific antigen or its analogue or
CC combinations. The method is useful for promoting nerve regeneration and
CC preventing neuronal degeneration in central/peripheral nervous system
CC from injury/disease, where the injury is spinal cord injury, blunt
CC trauma, penetrating trauma, hemorrhagic stroke, ischaemic stroke or

CC damages caused by surgery such as tumour excision. The disease is not an
CC autoimmune disease or neoplasm. The disease results in a degenerative
CC process occurring in either gray or white matter or both. The disease is
CC diabetic neuropathy, senile dementia, Alzheimer's disease, Parkinson's
CC disease, facial nerve (Bell's) palsy, glaucoma, Huntington's chorea,
CC amyotrophic lateral sclerosis, non-arteritic optic neuropathy, and
CC vitamin deficiency, intervertebral disc herniation, prion diseases such
CC as Creutzfeldt-Jakob disease, carpal tunnel syndrome, peripheral
CC neuropathies associated with various diseases, including but not limited
CC to uremia, porphyria, hypoglycemia, Sjorgren Larsson syndrome, acute
CC sensory neuropathy, chronic ataxic neuropathy, biliary cirrhosis, primary
CC amyloidosis, obstructive lung diseases, acromegaly, malabsorption
CC syndromes, polycythemia vera, immunoglobulin (Ig)A- and IgG gamma-
CC pathies, complications of various drugs (e.g., metronidazole) and toxins
CC (e.g., alcohol or organophosphates), Charcot-Marie-Tooth disease, ataxia
CC telangectasia, Friedreich's ataxia, amyloid polyneuropathies,
CC adrenomyeloneuropathy, Giant axonal neuropathy, Refsum's disease, Fabry's
CC disease, or lipoproteinemia. The present sequence represents the human
CC neurotransmitter receptor protein Nogo-A, an example of NS-specific
CC antigen
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 5; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171
```

```
Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 15
ABR59667
ID ABR59667 standard; protein; 1192 AA.
XX
AC ABR59667;
XX

DT 22-JUL-2003 (first entry)
XX
DE Human NogoA protein.
XX
KW Human; Nogo receptor; NgR; CTS domain; neuroprotective; gene therapy;
KW axonal growth; central nervous system; CNS; Nogo; spinal cord injury;
KW cranial trauma; cerebral trauma; spinal trauma; stroke; Krabbe's disease;
KW demyelinating disease; multiple sclerosis; monophasic demyelination;
KW encephalomyelitis; multifocal leukoencephalopathy; panencephalitis.
XX
OS Homo sapiens.
XX
PN WO2003031462-A2.
XX
PD 17-APR-2003.
XX
PF 04-OCT-2002; 2002WO-US032007.
XX
PR 06-OCT-2001; 2001US-00972599.
XX
PA (UYYA) UNIV YALE.
XX
PI Strittmatter SM;
XX
DR WPI; 2003-393433/37.
DR N-PSDB; ACC81048.
XX
PT New human Nogo receptor polypeptides and nucleic acids, useful for
PT decreasing inhibition of axonal growth by a central nervous system
PT neuron, or in treating central nervous system disease, disorder or
PT injury, e.g. spinal cord injury.
XX
PS Disclosure; Page 131-135; 148pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a polypeptide
CC comprising amino acid residues 27-309 of a 473 amino acid sequence (P1,
CC human Nogo receptor (NgR) NTLRRCT domain), or residues 27-309 of P1 with
CC 1-20 conservative amino acid substitutions, and less than a complete CTS
CC domain, provided that a partial CTS domain, if present, consists of no
CC more than the first 39 consecutive residues. The nucleic acid of the
CC invention has neuroprotective activity. The polynucleotide may have a use
CC in gene therapy. The nucleic acid is useful for decreasing inhibition of
CC axonal growth by a central nervous system (CNS) neuron. The NgR
CC polypeptide or an agent inhibits the binding of Nogo to NgR or NgR-
CC dependent signal transduction in the central nervous system neuron may be
CC used in treating central nervous system disease, disorder or injury, e.g.
CC spinal cord injury. Expression of an NgR protein may be associated with
CC inhibition of axonal regeneration following cranial, cerebral or spinal
CC trauma, stroke or a demyelinating disease, such as multiple sclerosis,
CC monophasic demyelination, encephalomyelitis, multifocal
CC leukoencephalopathy, panencephalitis, or Krabbe's disease. The present
CC sequence is used in the exemplification of the invention
XX
SQ Sequence 1192 AA;

Query Match 95.0%; Score 1756.9; DB 6; Length 1192;
Best Local Similarity 30.3%; Pred. No. 1.3e-60;

```
  Matches  361;  Conservative    0;  Mismatches    0;  Indels  831;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840
```

```
Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                     ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

Search completed: September 29, 2004, 18:51:25
Job time : 97.6283 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:48:38 ; Search time 23.5326 Seconds
(without alignments)
791.964 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	927	50.1	199	2	US-08-700-607-1	Sequence 1, Appli
2	827.7	44.7	776	2	US-08-700-607-5	Sequence 5, Appli
3	789.8	42.7	356	2	US-08-700-607-6	Sequence 6, Appli
4	684.2	37.0	208	2	US-08-700-607-7	Sequence 7, Appli
5	666	36.0	267	2	US-08-700-607-8	Sequence 8, Appli
6	541.9	29.3	168	4	US-09-149-476-563	Sequence 563, App
7	538.3	29.1	241	2	US-08-700-607-3	Sequence 3, Appli
8	348.9	18.9	75	4	US-09-621-976-4600	Sequence 4600, Ap
9	348.9	18.9	75	4	US-09-621-976-4601	Sequence 4601, Ap
10	285	15.4	92	4	US-09-149-476-411	Sequence 411, App
11	259.7	14.0	1248	2	US-09-080-897-2	Sequence 2, Appli

12	259.7	14.0	1248	3	US-09-323-735-2	Sequence 2, Appli
13	258.2	14.0	1315	3	US-08-899-595-3	Sequence 3, Appli
14	256.5	13.9	1255	2	US-09-080-897-4	Sequence 4, Appli
15	256.5	13.9	1255	3	US-08-899-595-1	Sequence 1, Appli
16	256.5	13.9	1255	3	US-09-323-735-4	Sequence 4, Appli
17	246.1	13.3	8991	4	US-08-714-741-32	Sequence 32, Appl
18	245.8	13.3	1070	3	US-08-922-635-22	Sequence 22, Appl
19	245.8	13.3	1504	4	US-09-364-206-2	Sequence 2, Appli
20	243.6	13.2	1719	2	US-08-459-568-4	Sequence 4, Appli
21	243.6	13.2	1719	2	US-08-399-411-4	Sequence 4, Appli
22	243.6	13.2	1719	3	US-08-516-859A-4	Sequence 4, Appli
23	243.6	13.2	1719	4	US-09-586-472-4	Sequence 4, Appli
24	243.6	13.2	1719	4	US-09-528-706-4	Sequence 4, Appli
25	241.5	13.1	933	3	US-08-764-870-14	Sequence 14, Appl
26	241.5	13.1	933	3	US-08-980-115-14	Sequence 14, Appl
27	238.7	12.9	1706	2	US-08-459-568-2	Sequence 2, Appli
28	238.7	12.9	1706	2	US-08-399-411-2	Sequence 2, Appli
29	238.7	12.9	1706	3	US-08-516-859A-2	Sequence 2, Appli
30	238.7	12.9	1706	4	US-09-586-472-2	Sequence 2, Appli
31	238.7	12.9	1706	4	US-09-528-706-2	Sequence 2, Appli
32	234.1	12.7	1326	4	US-09-688-188B-15	Sequence 15, Appl
33	234.1	12.7	1326	4	US-09-291-417D-15	Sequence 15, Appl
34	223.9	12.1	1266	4	US-08-506-296B-4	Sequence 4, Appli
35	223.9	12.1	2972	4	US-09-579-181-2	Sequence 2, Appli
36	223.9	12.1	3118	4	US-09-579-181-1	Sequence 1, Appli
37	223.4	12.1	1665	4	US-09-858-664A-2	Sequence 2, Appli
38	223.4	12.1	1665	4	US-10-274-978-2	Sequence 2, Appli
39	223.3	12.1	1274	4	US-09-095-443-2	Sequence 2, Appli
40	221.9	12.0	1601	4	US-09-345-473E-40	Sequence 40, Appl
41	221.7	12.0	559	4	US-10-116-370-2	Sequence 2, Appli
42	220.8	11.9	1185	3	US-09-041-886-23	Sequence 23, Appl
43	220.5	11.9	2205	1	US-08-093-453B-2	Sequence 2, Appli
44	219.2	11.8	905	2	US-08-574-959A-9	Sequence 9, Appli
45	219.2	11.8	905	3	US-09-357-014-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-08-700-607-1
; Sequence 1, Application US/08700607
; Patent No. 5858708
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
```

```
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 1.5
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/700,607
;       FILING DATE:  Filed Herewith
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J.
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0114 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-855-0555
;       TELEFAX:  415-845-4166
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  199 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     IMMEDIATE SOURCE:
;       LIBRARY:
;       CLONE:  Consensus
US-08-700-607-1

  Query Match             50.1%;  Score 927;  DB 2;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-34;
  Matches  188;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        174 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 71

Qy        234 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVL 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVL 131

Qy        294 MWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKI 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 MWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKI 191

Qy        354 PGLKRKAE 361
              ||||||||
Db        192 PGLKRKAE 199


RESULT 2
US-08-700-607-5
; Sequence 5, Application US/08700607
; Patent No. 5858708
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
```

```
;    TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,607
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0114 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  776 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  307307
US-08-700-607-5

  Query Match             44.7%;  Score 827.7;  DB 2;  Length 776;
  Best Local Similarity   27.1%;  Pred. No. 2e-28;
  Matches  196;  Conservative   55;  Mismatches   89;  Indels  383;  Gaps   17;

Qy          6 QSPL------------------------------------------------------- 9
              |||:
Db         70 QSPVAMETASTGVAGVSSAMDHTFSTTSKDGEGSCYTSLISDICYPPQEDSTYFTGILQK 129

Qy         10 ----VSSSDSPPR---PQP---------------------------------------- 21
                  |: |:||     | |
Db        130 ENGHVTISESPEELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKILADPLDQM 189

Qy         22 ---AFKYQFVREPEDEEEEEEEEEDEDEDLE---------------------------- 50
                 |:||  :  ||: : :|:   | ||:||:
Db        190 KAEAYKYIDITRPEEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKI 249

Qy         51 -----------------------------------ELE-------------------- 53
                                                  |:|
Db        250 IKDHLLEESTFAPYIDDLSEEQRRAPQITTPVKITLTEIEPSVETTTQEKTPEKQDICLK 309

Qy         54 ------------------------------------------------VLERKPAAGLS-- 64
```

```
                                                          :   | | |||
Db        310 PSPDTVPTVTVSEPEDDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEAKGLSYE 369

Qy         65 ----AAPV---------------PTAPAAGAPL-------------------------- 78
                 ||               || |   :||
Db        370 TAENPRPVGQLADRPEVKARSGPPTIP---SPLDHEASSAESGDSEIELVSEDPMAAEDA 426

Qy         79 -----MDFGNDFVPPAPRGPLPAAPPV-----------------------------APE 103
                : ||:   ||    | ||:| :                                :|:
Db        427 LPSGYVSFGHVGGPP----PSPASPSIQYSILREEREAELDSELIIESCDASSASEESPK 482

Qy        104 RQ---PSWDPSPVS-----------------------------STVPAPSPLSAAAVSP 130
               |:   |   || :                           || | | |
Db        483 REQDSPPMKPSALDAIREETGVRAEERAPSRRGLAEPGSFLDYPSTEPQPGP-------- 534

Qy        131 SKLPEDD-----EPPARPPPPPPASVSPQAEPVWTPPAP-APAAPP------STSVVDLL 178
               :||  |     | |  |  |   | | |:      | |  | |||          :|||
Db        535 -ELPPGDGALEPETPMLPRKPEEDSSSNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLL 593

Qy        179 YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDE 238
               ||||||:||:|||: | || ||| ||:||| ||:||| || |||||||| |:||:||:||
Db        594 YWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDE 653

Qy        239 GHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFT 298
               ||||:|||| |: :|:| :|||::    :|| |:|||||||| ||||||||||||: |
Db        654 GHPFKAYLELEITLSQEQIQKYTDCLQFYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLT 713

Qy        299 YVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKR 358
               |||||||||||::|::|:|::||:| :|||||| ||||   ::   :|||||||| ||
Db        714 YVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKR 773

Qy        359 KAE 361
               ||
Db        774 HAE 776
```

```
RESULT 3
US-08-700-607-6
; Sequence 6, Application US/08700607
; Patent No. 5858708
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
```

```
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 1.5
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/700,607
;       FILING DATE:  Filed Herewith
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J.
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0114 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-855-0555
;       TELEFAX:  415-845-4166
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  356 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     IMMEDIATE SOURCE:
;       LIBRARY:  GenBank
;       CLONE:  307309
US-08-700-607-6

  Query Match             42.7%;  Score 789.8;  DB 2;  Length 356;
  Best Local Similarity   42.6%;  Pred. No. 2.2e-27;
  Matches  169;  Conservative   42;  Mismatches   74;  Indels  112;  Gaps   10;

Qy         16 PPRP-QPAFKYQFVREPEDEE----------------EEEEEEEEDEDEDLEELEVLERK 58
              || |  |: :|  :||  : |                 ||  : |:|
Db         21 PPSPASPSIQYSILREEREAELDSELIIESCDASSASEESPKREQD-------------- 66

Qy         59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAP---------------E 103
                                                 :||: |                |
Db         67 ------------------------------------SPPMKPSALDAIREETGVRAEE 88

Qy        104 RQPS----WDPS---PVSSTVPAPSPLSAAAVSPSKLPEDD-----EPPARPPPPPASV 151
              | ||     :|       || | | |         :||  |     | |  |  |   |
Db         89 RAPSRRGLAEPGSFLDYPSTEPQPGP---------ELPPGDGALEPETPMLPRKPEEDSS 139

Qy        152 SPQAEPVWTPPAP-APAAPP------STSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS 204
              | |:      | |  | |||          :|||||||||:||:|||: | || ||| ||
Db        140 SNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFS 199

Qy        205 IVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA 264
              :||| ||:||| || |||||||| |:||:||:||||||:|||| |: :|:| :|||::
Db        200 VVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTDCL 259

Qy        265 LGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIY 324
                :|| |:|||||||||| |||||||||||||: |||||||||||||::|::|:|::||:|
Db        260 QFYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVY 319

Qy        325 ERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
               :|||||| ||||   ::   :|||||||| || ||
Db        320 VKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKRHAE 356
```

RESULT 4
US-08-700-607-7
; Sequence 7, Application US/08700607
; Patent No. 5858708
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: TWO NOVEL HUMAN NSP-LIKE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,607
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 307311
US-08-700-607-7

Query Match 37.0%; Score 684.2; DB 2; Length 208;
Best Local Similarity 62.4%; Pred. No. 3.9e-23;
Matches 131; Conservative 35; Mismatches 36; Indels 8; Gaps 1;

```
Qy        152 SPQAEPVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAY 211
              | : : ||        :   :  :|||||||||:||:|||: | || ||| ||:||| ||
Db          7 STKMDCVW--------SNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAY 58

Qy        212 IALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCT 271
              :||| || |||||||| |:||:||:||||||:|||| |: :|:| :|||::    :|| |
Db         59 LALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTDCLQFYVNST 118
```

```
Qy        272 IKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQI 331
              :|||||||||  |||||||||||||: |||||||||||||::|::|:|::||:| :|||||
Db        119 LKELRRLFLVQDLVDSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQI 178

Qy        332 DHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              | ||||   ::   :|||||||| || ||
Db        179 DQYLGLVRTHINAVVAKIQAKIPGAKRHAE 208


RESULT 5
US-08-700-607-8
; Sequence 8, Application US/08700607
; Patent No. 5858708
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,607
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0114 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  281046
US-08-700-607-8

  Query Match             36.0%;  Score 666;  DB 2;  Length 267;
  Best Local Similarity   66.1%;  Pred. No. 4e-22;
  Matches  123;  Conservative   34;  Mismatches   29;  Indels    0;  Gaps    0;
```

```
Qy        172 TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQ 231
              :  :|||||||||:||:|||: | || ||| ||:||| ||:||| || |||||||| |:|
Db         10 SQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQ 69

Qy        232 AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFA 291
              |:||:||||||:|||| |: :|:| :|||::    :|| |:|||||||| ||||||||
Db         70 AVQKTDEGHPFKAYLELEITLSQEQIQKYTDCLQLYVNSTLKELRRLFLVQDLVDSLKFA 129

Qy        292 VLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA 351
              ||||: |||||||||||||::|::|:|::||:| :|||||:| ||||   ::   :|||||
Db        130 VLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQVDQYLGLVRTHINTVVAKIQA 189

Qy        352 KIPGLK 357
              |||| :
Db        190 KIPGAR 195
```

```
RESULT 6
US-09-149-476-563
; Sequence 563, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
```

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 29.3%; Score 541.9; DB 4; Length 168;
Best Local Similarity 60.5%; Pred. No. 5.4e-17;
Matches 101; Conservative 34; Mismatches 31; Indels 1; Gaps 1;

```
Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEE 255
              :||||  ||:::|| :|: |||||||||||||| ||||:|||:||||||:|||: :: :| |
Db          1 MLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKAYLDVDITLSSE 60
```

```
Qy        256 LVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALI 315
                  | |:|: |:|  :| : |||||:||||||| || ||: |||||:|||:|||||| :
Db         61 AFHNYMNAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFNGITLLILAEL 120

Qy        316 SLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-KRKAE 361
               :||||::||::: |||||:|:|    |  : |||||:||: |:|||
Db        121 LIFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKKAE 167
```

```
RESULT 7
US-08-700-607-3
; Sequence 3, Application US/08700607
; Patent No. 5858708
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    TITLE OF INVENTION:  TWO NOVEL HUMAN NSP-LIKE PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  U.S.
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,607
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0114 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  241 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  THP1NOB01
;      CLONE:  31870
US-08-700-607-3

  Query Match             29.1%;  Score 538.3;  DB 2;  Length 241;
  Best Local Similarity   30.9%;  Pred. No. 1.6e-16;
```

```
  Matches  112;  Conservative   38;  Mismatches   46;  Indels  167;  Gaps    4;

Qy          6 QSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPAAGLSA 65
              ||  :|||                                                    |
Db          9 QSHSISSS----------------------------------------------------SF 18

Qy         66 APVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAPSPLSA 125
                 |:||  |
Db         19 GAEPSAPGGGG------------------------------------------------- 29

Qy        126 AAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPP---------STSVVD 176
                                                 :| | |         | :| |
Db         30 -------------------------------------SPGACPALGTKSCSSSCAVHD 50

Qy        177 LLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKS 236
              |: |||:|||| ||| :| :|||| ||::|| :|: |||||||||||||| ||||:|||
Db         51 LIXWRDVKKTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAVQKS 110

Qy        237 DEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWV 296
              :|||||:|||: :: :| |    | |:|: |:|  :| : |||||:||||||| || ||:
Db        111 EEGHPFKAYLDVDITLSSEAFHNYMNAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWL 170

Qy        297 FTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL 356
               |||||:|||:|||||| : : |||::| ::                    |: :| |
Db        171 MTYVGAVFNGITLLILAELLIXSVPIVYXKY--------------------KVPSKTPWN 210

Qy        357 KRK 359
              ::|
Db        211 RQK 213


RESULT 8
US-09-621-976-4600
; Sequence 4600, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 4600
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 58
;   OTHER INFORMATION: Xaa = His,Pro
;   NAME/KEY: UNSURE
;   LOCATION: 28
;   OTHER INFORMATION: Xaa = Met,Val
;   NAME/KEY: UNSURE
```

```
;   LOCATION: 19
;   OTHER INFORMATION: Xaa = Pro,Gln
;   NAME/KEY: UNSURE
;   LOCATION: 53
;   OTHER INFORMATION: Xaa = Ser,Tyr
US-09-621-976-4600

  Query Match             18.9%;  Score 348.9;  DB 4;  Length 75;
  Best Local Similarity   82.6%;  Pred. No. 4.4e-09;
  Matches   71;  Conservative    0;  Mismatches    4;  Indels   11;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              |||||||||||||||||| |||||||| |||||||           ||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRXQPAFKYQFXREPEDEE----------EDLEELEVLERKPA 49

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFV 86
              ||| |||| ||||||||||||||||||
Db         50 AGLXAAPVXTAPAAGAPLMDFGNDFV 75


RESULT 9
US-09-621-976-4601
; Sequence 4601, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 4601
;   LENGTH: 75
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 58
;   OTHER INFORMATION: Xaa = His,Pro
;   NAME/KEY: UNSURE
;   LOCATION: 28
;   OTHER INFORMATION: Xaa = Met,Val
;   NAME/KEY: UNSURE
;   LOCATION: 19
;   OTHER INFORMATION: Xaa = Pro,Gln
;   NAME/KEY: UNSURE
;   LOCATION: 53
;   OTHER INFORMATION: Xaa = Ser,Tyr
US-09-621-976-4601

  Query Match             18.9%;  Score 348.9;  DB 4;  Length 75;
  Best Local Similarity   82.6%;  Pred. No. 4.4e-09;
  Matches   71;  Conservative    0;  Mismatches    4;  Indels   11;  Gaps    1;
```

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||| |||||||| |||||||           ||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRXQPAFKYQFXREPEDEE-----------EDLEELEVLERKPA 49

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFV 86
              ||| |||| ||||||||||||||||||
Db         50 AGLXAAPVXTAPAAGAPLMDFGNDFV 75
```

RESULT 10
US-09-149-476-411
; Sequence 411, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23

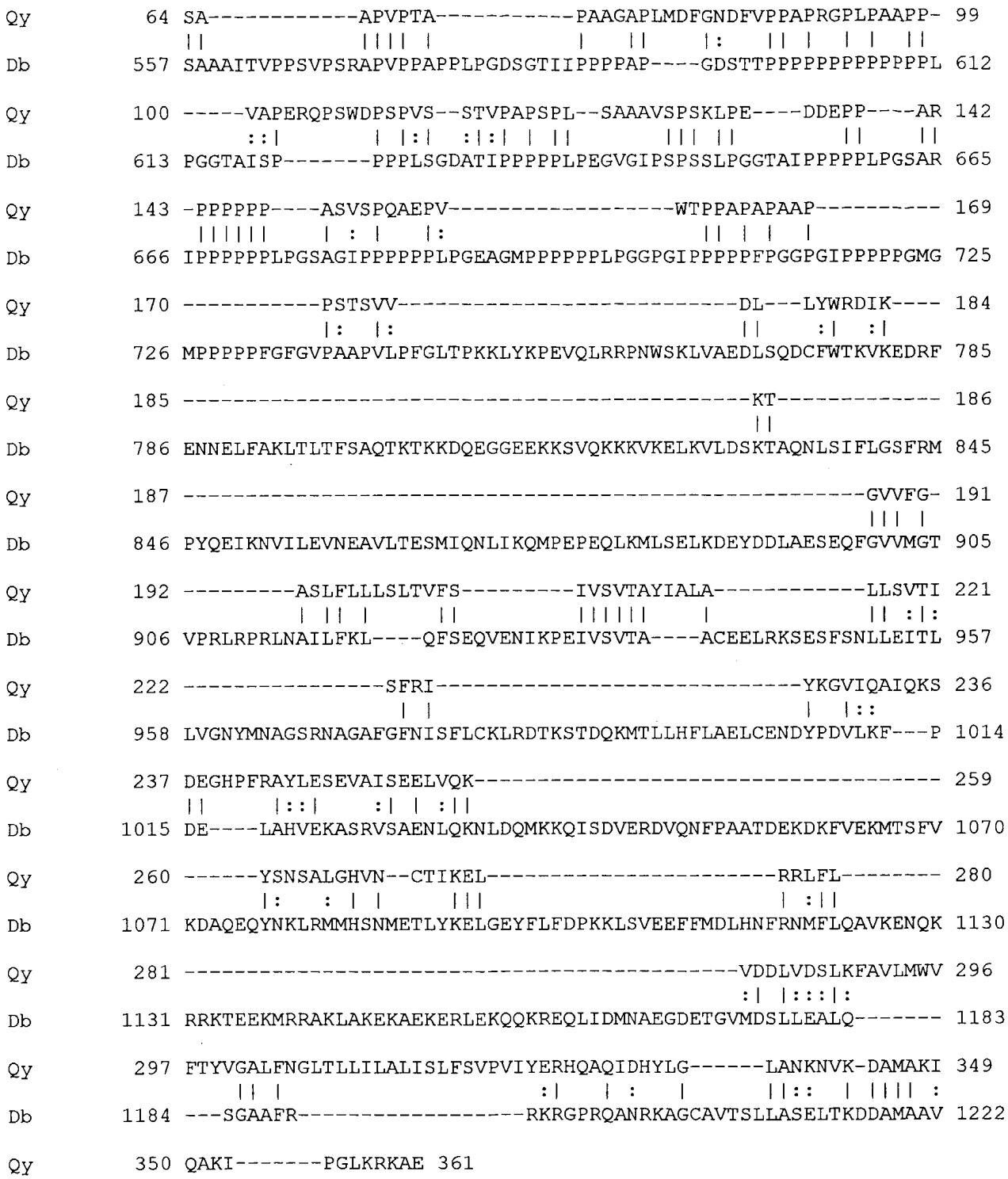
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

  Query Match   15.4%; Score 285; DB 4; Length 92;
  Best Local Similarity 57.1%; Pred. No. 4.5e-06;
  Matches   52; Conservative   20; Mismatches   19; Indels   0; Gaps   0;

```
Qy        262 NSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVP 321
              |:|: |:|  :| : |||||:||||||| || ||: |||||:|||:|||||| : :||||
Db          2 NAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFNGITLLILAELLIFSVP 61

Qy        322 VIYERHQAQIDHYLGLANKNVKDAMAKIQAK 352
              ::||::: |||||:|:|    |  : || :|
Db         62 IVYEKYKTQIDHYVGIARDQTKSIVEKIPSK 92
```

RESULT 11
US-09-080-897-2
; Sequence 2, Application US/09080897

```
; Patent No. 5985574
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,897
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1248 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-080-897-2

  Query Match             14.0%;  Score 259.7;  DB 2;  Length 1248;
  Best Local Similarity   15.6%;  Pred. No. 0.0088;
  Matches  143;  Conservative   49;  Mismatches  114;  Indels  613;  Gaps   40;

Qy          1 MEDLD----------------------QSPLVSSSDSPPRPQ----------------- 20
                |:| :                        |  |:  :|   |||
Db        378 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVRNDYEARPQYYKLIEECISQIVLHKN 437

Qy         21 ---PAF---------------------------------------------------- 23
                  | |
Db        438 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 497

Qy         24 -----KYQFVREPEDEEEEEEEEEDEDEDLE---------------ELEVLERKPAAGL 63
                   | | ::  :|    |:::   | :|||               |||   :|  | |
Db        498 SDFEQKLQDLQGEKDALHSEKQQIATEKQDLEAEVSQLTGEVAKLTKELED-AKKEMASL 556
```



RESULT 12
US-09-323-735-2
; Sequence 2, Application US/09323735
; Patent No. 6197932

```
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/323,735
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/080,897
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1248 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-323-735-2

  Query Match             14.0%;  Score 259.7;  DB 3;  Length 1248;
  Best Local Similarity   15.6%;  Pred. No. 0.0088;
  Matches  143;  Conservative   49;  Mismatches  114;  Indels  613;  Gaps   40;

Qy          1 MEDLD----------------------QSPLVSSSDSPPRPQ----------------- 20
                |:| :                      |  |:  :|   |||
Db        378 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVRNDYEARPQYYKLIEECISQIVLHKN 437

Qy         21 ---PAF---------------------------------------------------- 23
                 | |
Db        438 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 497

Qy         24 -----KYQFVREPEDEEEEEEEEEEDEDEDLE---------------ELEVLERKPAAGL 63
                  | | ::  :|    |:::   | :|||               |||   :|  | |
```

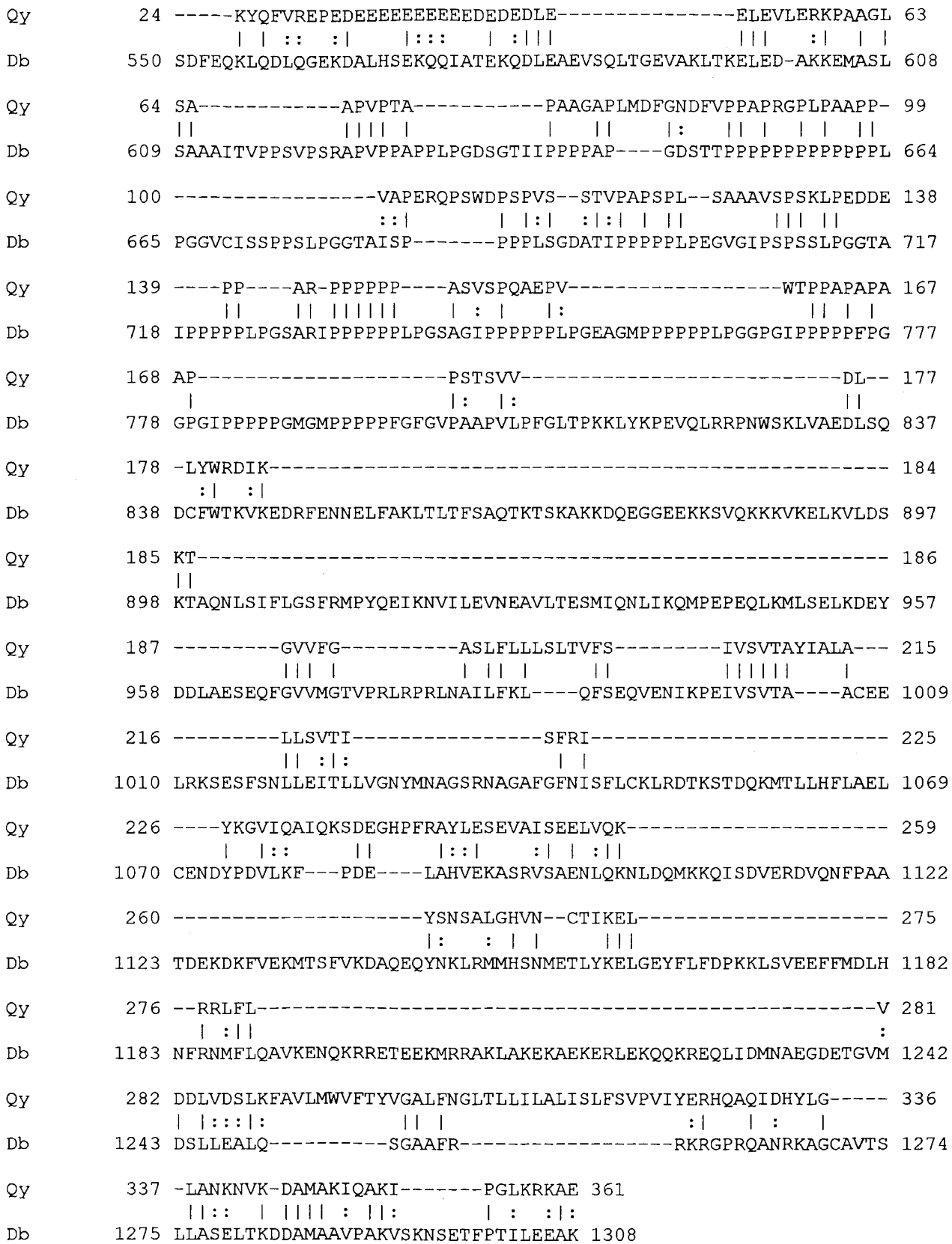
```
Db        498 SDFEQKLQDLQGEKDALHSEKQQIATEKQDLEAEVSQLTGEVAKLTKELED-AKKEMASL 556

Qy         64 SA------------APVPTA-----------PAAGAPLMDFGNDFVPPAPRGPLPAAPP- 99
              ||            |||| |           |   ||    |:   || |  | |  ||
Db        557 SAAAITVPPSVPSRAPVPPAPPLPGDSGTIIPPPPAP----GDSTTPPPPPPPPPPPPL 612

Qy        100 -----VAPERQPSWDPSPVS--STVPAPSPL--SAAAVSPSKLPE----DDEPP----AR 142
                   ::|       | |:|  :|:| | ||       ||| ||         ||    ||
Db        613 PGGTAISP-------PPPLSGDATIPPPPPLPEGVGIPSPSSLPGGTAIPPPPPLPGSAR 665

Qy        143 -PPPPPP----ASVSPQAEPV-----------------WTPPAPAPAAP---------- 169
                |||||     | : |   |:                  || | |  |
Db        666 IPPPPPPLPGSAGIPPPPPPLPGEAGMPPPPPPLPGGPGIPPPPPFPGGPGIPPPPPGMG 725

Qy        170 -----------PSTSVV-------------------------DL---LYWRDIK---- 184
                         |:  |:                         ||    :|  :|
Db        726 MPPPPPFGFGVPAAPVLPFGLTPKKLYKPEVQLRRPNWSKLVAEDLSQDCFWTKVKEDRF 785

Qy        185 ----------------------------------------------KT---------- 186
                                                           ||
Db        786 ENNELFAKLTLTFSAQTKTKKDQEGGEEKKSVQKKKVKELKVLDSKTAQNLSIFLGSFRM 845

Qy        187 ------------------------------------------------------GVVFG- 191
                                                                   ||| |
Db        846 PYQEIKNVILEVNEAVLTESMIQNLIKQMPEPEQLKMLSELKDEYDDLAESEQFGVVMGT 905

Qy        192 ---------ASLFLLLSLTVFS---------IVSVTAYIALA----------LLSVTI 221
                      | || |     ||         ||||||    |           || :|:
Db        906 VPRLRPRLNAILFKL----QFSEQVENIKPEIVSVTA----ACEELRKSESFSNLLEITL 957

Qy        222 ----------------SFRI----------------------------YKGVIQAIQKS 236
                           | |                             |  |::
Db        958 LVGNYMNAGSRNAGAFGFNISFLCKLRDTKSTDQKMTLLHFLAELCENDYPDVLKF---P 1014

Qy        237 DEGHPFRAYLESEVAISEELVQK---------------------------------- 259
              ||     |::|    :| | :||
Db       1015 DE----LAHVEKASRVSAENLQKNLDQMKKQISDVERDVQNFPAATDEKDKFVEKMTSFV 1070

Qy        260 ------YSNSALGHVN--CTIKEL----------------------RRLFL-------- 280
                    |:   : | |     |||                      | :||
Db       1071 KDAQEQYNKLRMMHSNMETLYKELGEYFLFDPKKLSVEEFFMDLHNFRNMFLQAVKENQK 1130

Qy        281 ------------------------------------------VDDLVDSLKFAVLMWV 296
                                                       :| |:::|:
Db       1131 RRKTEEKMRRAKLAKEKAEKERLEKQQKREQLIDMNAEGDETGVMDSLLEALQ------- 1183

Qy        297 FTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG------LANKNVK-DAMAKI 349
                 || |                   :|   | :   |      ||::  | |||| :
Db       1184 ---SGAAFR----------------RKRGPRQANRKAGCAVTSLLASELTKDDAMAAV 1222

Qy        350 QAKI-------PGLKRKAE 361
               ||:       | :  :|:
Db       1223 PAKVSKNSETFPTILEEAK 1241
```

RESULT 13
US-08-899-595-3

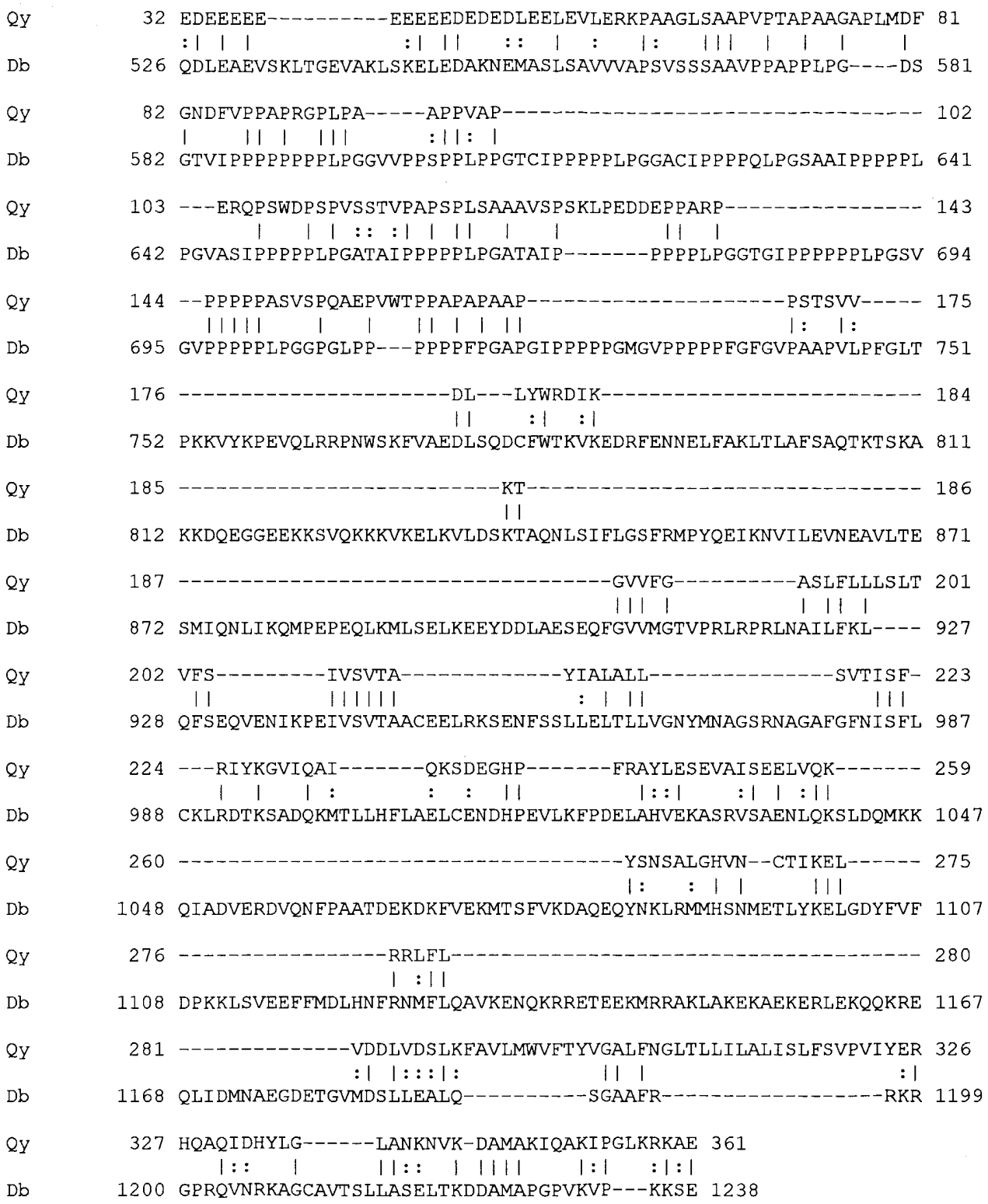
; Sequence 3, Application US/08899595
; Patent No. 6111072
; GENERAL INFORMATION:
; APPLICANT: Narumiya, Shuh
; APPLICANT: Takahashi, No. 6111072uaki
; TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE
; TITLE OF INVENTION: ENCODING SAME
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,595
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-242701
; FILING DATE: 26-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-90170
; FILING DATE: 25-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Stephen A. Bent
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 049441/0112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-899-595-3

Query Match 14.0%; Score 258.2; DB 3; Length 1315;
Best Local Similarity 15.3%; Pred. No. 0.011;
Matches 143; Conservative 49; Mismatches 114; Indels 628; Gaps 40;

```
Qy          1 MEDLD----------------------QSPLVSSSDSPPRPQ----------------- 20
                |:| :                      |  |:  :|   |||
Db        430 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVRNDYEARPQYYKLIEECISQIVLHKN 489

Qy         21 ---PAF---------------------------------------------------- 23
                  | |
Db        490 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 549
```



```
RESULT 14
US-09-080-897-4
; Sequence 4, Application US/09080897
; Patent No. 5985574
;  GENERAL INFORMATION:
;    APPLICANT:  King, Mary-Claire
;    APPLICANT:  Lynch, Eric D.
;    APPLICANT:  Lee, Ming
;    APPLICANT:  Morrow, Jan E.
;    APPLICANT:  Welcsh, Piri L.
;    APPLICANT:  Leon, Pedro E.
;    TITLE OF INVENTION:  Modulators of Actin
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
;      STREET:  75 DENISE DRIVE
;      CITY:  HILLSBOROUGH
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/080,897
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OSMAN, RICHARD A
;      REGISTRATION NUMBER:  36,627
;      REFERENCE/DOCKET NUMBER:  UW97-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 343-4341
;      TELEFAX:  (650) 343-4342
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-080-897-4

  Query Match             13.9%;  Score 256.5;  DB 2;  Length 1255;
  Best Local Similarity   14.7%;  Pred. No. 0.012;
  Matches  130;  Conservative   42;  Mismatches  135;  Indels  575;  Gaps   31;

Qy          6 QSPLVSSSDSPPRPQ--------------------PAF---------------------- 23
              |  |:  :|   |||                    | |
Db        406 QHLLLVRNDYEARPQYYKLIEECVSQIVLHKNGTDPDFKCRHLQIDIERLVDQMIDKTKV 465

Qy         24 ------------------------------------------------------KYQFVREP 31
                                                                    | |   :
Db        466 EKSEAKATELEKKLDSELTARHELQVEMKKMENDFEQKLQDLQGEKDALDSEKQQITAQK 525
```



RESULT 15
US-08-899-595-1
; Sequence 1, Application US/08899595

```
; Patent No. 6111072
;  GENERAL INFORMATION:
;    APPLICANT:  Narumiya, Shuh
;    APPLICANT:  Takahashi, No. 6111072uaki
;    TITLE OF INVENTION:  RHO TARGET PROTEIN HUMAN MDIA AND GENE
;    TITLE OF INVENTION:  ENCODING SAME
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/899,595
;      FILING DATE:  24-JUL-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-242701
;      FILING DATE:  26-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 9-90170
;      FILING DATE:  25-MAR-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stephen A. Bent
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  049441/0112
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1255 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-899-595-1

  Query Match             13.9%;  Score 256.5;  DB 3;  Length 1255;
  Best Local Similarity   14.7%;  Pred. No. 0.012;
  Matches  130;  Conservative   42;  Mismatches  135;  Indels  575;  Gaps   31;

Qy          6 QSPLVSSSDSPPRPQ--------------------PAF---------------------- 23
              |  |:  :|   |||                     | |
Db        406 QHLLLVRNDYEARPQYYKLIEECVSQIVLHKNGTDPDFKCRHLQIDIERLVDQMIDKTKV 465

Qy         24 ------------------------------------------------------KYQFVREP 31
                                                                    | |   :
Db        466 EKSEAKATELEKKLDSELTARHELQVEMKKMENDFEQKLQDLQGEKDALDSEKQQITAQK 525
```

```
Qy         32 EDEEEEE----------EEEEEDEDEDLEELEVLERKPAAGLSAAPVPTAPAAGAPLMDF 81
              :| | |           :| ||   ::  |  :   |:   |||  |  |  |    |
Db        526 QDLEAEVSKLTGEVAKLSKELEDAKNEMASLSAVVVAPSVSSSAAVPPAPPLPG----DS 581

Qy         82 GNDFVPPAPRGPLPA-----APPVAP---------------------------------- 102
              |    || |  |||      :||: |
Db        582 GTVIPPPPPPPPLPGGVVPPSPPLPPGTCIPPPPPLPGGACIPPPPQLPGSAAIPPPPPL 641

Qy        103 ---ERQPSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPARP---------------- 143
                   |   | | :: :| | ||  |   |        ||  |
Db        642 PGVASIPPPPPLPGATAIPPPPPLPGATAIP-------PPPPLPGGTGIPPPPPPLPGSV 694

Qy        144 --PPPPPASVSPQAEPVWTPPAPAPAAP---------------------PSTSVV----- 175
                |||||    |   |   || | | ||                        |:  |:
Db        695 GVPPPPPLPGGPGLPP---PPPPFPGAPGIPPPPPGMGVPPPPPFGFGVPAAPVLPFGLT 751

Qy        176 ----------------------DL---LYWRDIK------------------------- 184
                                   ||    :|  :|
Db        752 PKKVYKPEVQLRRPNWSKFVAEDLSQDCFWTKVKEDRFENNELFAKLTLAFSAQTKTSKA 811

Qy        185 --------------------------KT-------------------------------- 186
                                        ||
Db        812 KKDQEGGEEKKSVQKKKVKELKVLDSKTAQNLSIFLGSFRMPYQEIKNVILEVNEAVLTE 871

Qy        187 ----------------------------------GVVFG----------ASLFLLLSLT 201
                                               ||| |          | || |
Db        872 SMIQNLIKQMPEPEQLKMLSELKEEYDDLAESEQFGVVMGTVPRLRPRLNAILFKL---- 927

Qy        202 VFS---------IVSVTA-------------YIALALL---------------SVTISF- 223
               ||         ||||||              : | ||                  |||
Db        928 QFSEQVENIKPEIVSVTAACEELRKSENFSSLLELTLLVGNYMNAGSRNAGAFGFNISFL 987

Qy        224 ---RIYKGVIQAI-------QKSDEGHP-------FRAYLESEVAISEELVQK------- 259
                  |  |   | :       :  :  ||         |::|    :| | :||
Db        988 CKLRDTKSADQKMTLLHFLAELCENDHPEVLKFPDELAHVEKASRVSAENLQKSLDQMKK 1047

Qy        260 ------------------------------------YSNSALGHVN--CTIKEL------ 275
                                                 |:   : | |     |||
Db       1048 QIADVERDVQNFPAATDEKDKFVEKMTSFVKDAQEQYNKLRMMHSNMETLYKELGDYFVF 1107

Qy        276 -----------------RRLFL---------------------------------------- 280
                             | :||
Db       1108 DPKKLSVEEFFMDLHNFRNMFLQAVKENQKRRETEEKMRRAKLAKEKAEKERLEKQQKRE 1167

Qy        281 --------------VDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYER 326
                          :| |:::|:           || |                    :|
Db       1168 QLIDMNAEGDETGVMDSLLEALQ----------SGAAFR-----------------RKR 1199

Qy        327 HQAQIDHYLG------LANKNVK-DAMAKIQAKIPGLKRKAE 361
                |::   |      ||::  | ||||    |:|   :|:|
Db       1200 GPRQVNRKAGCAVTSLLASELTKDDAMAPGPVKVP---KKSE 1238
```

Search completed: September 29, 2004, 18:56:38
Job time : 27.0326 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:43:17 ; Search time 26.5368 Seconds
(without alignments)
1308.568 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	827.7	44.7	776	2	A46583	neuroendocrine-spe
2	684.2	37.0	208	2	I60904	neuroendocrine-spe
3	666	36.0	267	2	A60021	tropomyosin-relate
4	454.3	24.6	2484	2	T26216	hypothetical prote
5	447	24.2	2607	2	T26215	hypothetical prote
6	353.5	19.1	222	2	T26213	hypothetical prote
7	275.1	14.9	1173	2	T31421	C-terminal domain-
8	266	14.4	1206	2	S24407	formin isoform IV
9	266	14.4	1468	2	S11515	formin - mouse
10	264.1	14.3	1058	2	T13286	cappuccino gene pr
11	262.4	14.2	1611	2	T38236	hypothetical prote
12	260.9	14.1	760	2	F86387	probable Pto kinas
13	257.1	13.9	3511	2	A59295	unconventional myo

14	256.5	13.9	1255	2	T31065	diaphanous protein
15	255.5	13.8	3530	2	A59266	unconventional myo
16	253.8	13.7	1132	2	A35098	MHC class III hist
17	253.6	13.7	1201	2	G86441	unknown protein [i
18	250.4	13.5	880	2	D89756	protein T23E7.2b [
19	250	13.5	980	2	S54986	regulatory protein
20	247.7	13.4	1015	2	JC6552	DNA topoisomerase
21	247.5	13.4	731	2	B86369	hypothetical prote
22	247.3	13.4	980	2	G75523	probable cell divi
23	247.2	13.4	1047	2	A55617	masquerade precurs
24	247	13.4	2090	2	S26058	probable transform
25	244.8	13.2	1127	2	T32404	hypothetical prote
26	244.8	13.2	1375	2	S48375	hypothetical prote
27	243.6	13.2	1721	1	I38902	retinoblastoma bin
28	243.4	13.2	3938	2	T42761	Bassoon protein -
29	243.3	13.2	708	2	D96711	hypothetical prote
30	243.3	13.2	786	2	A35466	progesterone recep
31	243	13.1	907	2	E96636	hypothetical prote
32	242.9	13.1	1634	2	T26517	hypothetical prote
33	242.6	13.1	716	2	T26998	hypothetical prote
34	241.5	13.1	933	1	QRHUP	progesterone recep
35	239.2	12.9	449	2	S16748	proline-rich prote
36	239.1	12.9	930	2	A25923	progesterone recep
37	238.7	12.9	1706	2	I84499	zinc finger protei
38	238.3	12.9	3149	1	QQBE8	BPLF1 protein - hu
39	237	12.8	1137	2	A86335	T20H2.9 protein -
40	236.9	12.8	710	2	D96728	hypothetical prote
41	236.7	12.8	2649	2	T51023	hypothetical prote
42	235.6	12.7	765	2	T49592	neurofilament trip
43	235.2	12.7	1953	2	S63244	BNI1 protein - yea
44	235.1	12.7	534	2	S21961	proline-rich prote
45	234.8	12.7	993	2	T17230	hypothetical prote

ALIGNMENTS

RESULT 1
A46583
neuroendocrine-specific protein, splice form A - human
N;Contains: neuroendocrine-specific protein, splice form B
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999
C;Accession: A46583; I60903
R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.
J. Biol. Chem. 268, 13439-13447, 1993
A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.
A;Reference number: A46583; MUID:93293865; PMID:7685762
A;Accession: A46583
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-776 <ROE1>
A;Cross-references: GB:L10333; NID:g307306; PIDN:AAA59950.1; PID:g307307
A;Accession: I60903

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 421-776 <ROE2>
A;Cross-references: GB:L10334; NID:g307308; PIDN:AAA59951.1; PID:g307309
C;Genetics:
A;Gene: GDB:RTN1; NSP
A;Cross-references: GDB:203968; OMIM:600865
A;Map position: 14q21-14q22

Query Match 44.7%; Score 827.7; DB 2; Length 776;
Best Local Similarity 27.1%; Pred. No. 3.3e-20;
Matches 196; Conservative 55; Mismatches 89; Indels 383; Gaps 17;

```
Qy          6 QSPL------------------------------------------------------------ 9
                |||:
Db         70 QSPVAMETASTGVAGVSSAMDHTFSTTSKDGEGSCYTSLISDICYPPQEDSTYFTGILQK 129

Qy         10 ----VSSSDSPPR---PQP---------------------------------------------- 21
                   |: |:||     | |
Db        130 ENGHVTISESPEELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKILADPLDQM 189

Qy         22 ---AFKYQFVREPEDEEEEEEEEEDEDEDLE---------------------------- 50
                  |:||  :  ||: : :|:   | ||:||:
Db        190 KAEAYKYIDITRPEEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKI 249

Qy         51 ------------------------------------ELE-------------------- 53
                                                   |:|
Db        250 IKDHLLEESTFAPYIDDLSEEQRRAPQITTPVKITLTEIEPSVETTTQEKTPEKQDICLK 309

Qy         54 ------------------------------------------------VLERKPAAGLS-- 64
                                                            :   | | |||
Db        310 PSPDTVPTVTVSEPEDDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEAKGLSYE 369

Qy         65 ----AAPV---------------PTAPAAGAPL--------------------------- 78
                    ||               || |   :||
Db        370 TAENPRPVGQLADRPEVKARSGPPTIP---SPLDHEASSAESGDSEIELVSEDPMAAEDA 426

Qy         79 -----MDFGNDFVPPAPRGPLPAAPPV-----------------------------APE 103
                   : ||:   ||    | ||:| :                            :|:
Db        427 LPSGYVSFGHVGGPP----PSPASPSIQYSILREEREAELDSELIIESCDASSASEESPK 482

Qy        104 RQ---PSWDPSPVS------------------------------STVPAPSPLSAAAVSP 130
                |:   |   || :                           || | | |
Db        483 REQDSPPMKPSALDAIREETGVRAEERAPSRRGLAEPGSFLDYPSTEPQPGP-------- 534

Qy        131 SKLPEDD-----EPPARPPPPPASVSPQAEPVWTPPAP-APAAPP------STSVVDLL 178
                :||  |     | |  |  |   | | |:      | |  | |||          :|||
Db        535 -ELPPGDGALEPETPMLPRKPEEDSSSNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLL 593

Qy        179 YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDE 238
              ||||||:||:|||: | || ||| ||:||| ||:||| || |||||||| |:||:||:||
Db        594 YWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDE 653

Qy        239 GHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFT 298
              ||||:|||| |: :|:| :|||::    :|| |:||||||||| ||||||||||||: |
Db        654 GHPFKAYLELEITLSQEQIQKYTDCLQFYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLT 713
```

```
Qy        299 YVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKR 358
              ||||||||||||::|::|:|::||:| :|||||| ||||   ::   :|||||||||| ||
Db        714 YVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKR 773

Qy        359 KAE 361
              ||
Db        774 HAE 776
```

RESULT 2
I60904
neuroendocrine-specific protein C - human
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 05-Nov-1999
C;Accession: I60904
R;Roebroek, A.J.; van de Velde, H.J.; Van Bokhoven, A.; Broers, J.L.; Ramaekers, F.C.; Van de Ven, W.J.
J. Biol. Chem. 268, 13439-13447, 1993
A;Title: Cloning and expression of alternative transcripts of a novel neuroendocrine-specific gene and identification of its 135-kDa translational product.
A;Reference number: A46583; MUID:93293865; PMID:7685762
A;Accession: I60904
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-208 <RES>
A;Cross-references: GB:L10335; NID:g307310; PIDN:AAA59952.1; PID:g307311
C;Genetics:
A;Gene: GDB:RTN1; NSP
A;Cross-references: GDB:203968; OMIM:600865
A;Map position: 14q21-14q22

```
  Query Match             37.0%;  Score 684.2;  DB 2;  Length 208;
  Best Local Similarity   62.4%;  Pred. No. 9.2e-17;
  Matches  131;  Conservative   35;  Mismatches   36;  Indels    8;  Gaps    1;

Qy        152 SPQAEPVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAY 211
              | : : ||        :   :  :||||||||||:||:|||: | || ||| ||:||| ||
Db          7 STKMDCVW--------SNWKSQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAY 58

Qy        212 IALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCT 271
              :||| || |||||||| |:||:||:||||||:|||| |: :|:| :|||::    :|| |
Db         59 LALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKYTDCLQFYVNST 118

Qy        272 IKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQI 331
              :|||||||| ||||||||||||: ||||||||||||::|::|:|::||:| :||||
Db        119 LKELRRLFLVQDLVDSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQI 178

Qy        332 DHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              | ||||   ::   :|||||||| || ||
Db        179 DQYLGLVRTHINAVVAKIQAKIPGAKRHAE 208
```

RESULT 3
A60021
tropomyosin-related protein, neuronal - rat
C;Species: Rattus norvegicus (Norway rat)

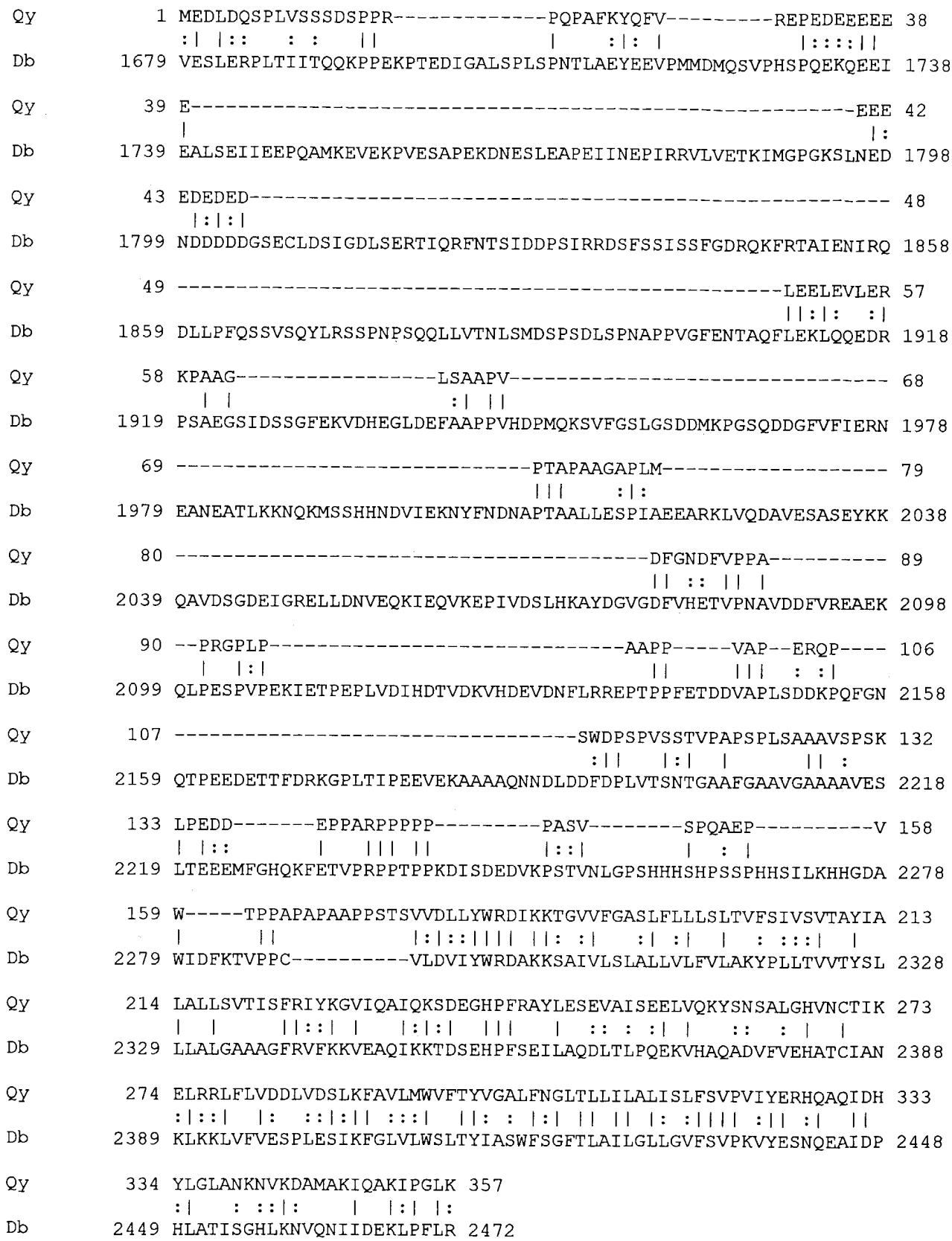
C;Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 05-Nov-1999
C;Accession: A60021
R;Wieczorek, D.F.; Hughes, S.R.
Brain Res. Mol. Brain Res. 10, 33-41, 1991
A;Title: Developmentally regulated cDNA expressed exclusively in neural tissue.
A;Reference number: A60021; MUID:91278684; PMID:1647480
A;Accession: A60021
A;Molecule type: mRNA
A;Residues: 1-267 <WIE>
A;Cross-references: EMBL:X52817; NID:g456549; PIDN:CAA37001.1; PID:g456550
C;Comment: This neuronal-specific mRNA was identified by hybridization to an alpha-tropomyosin probe but does not show homology in amino acid sequence.

  Query Match 36.0%; Score 666; DB 2; Length 267;
  Best Local Similarity 66.1%; Pred. No. 7.3e-16;
  Matches 123; Conservative 34; Mismatches 29; Indels 0; Gaps 0;

```
Qy        172 TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQ 231
              :  :|||||||||:||:|||: | || ||| ||:||| ||:||| || |||||||| |:|
Db         10 SQAIDLLYWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAAALSATISFRIYKSVLQ 69

Qy        232 AIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFA 291
              |:||:||||||:|||| |: :|:| :|||::    :|| |:|||||||| |||||||||
Db         70 AVQKTDEGHPFKAYLELEITLSQEQIQKYTDCLQLYVNSTLKELRRLFLVQDLVDSLKFA 129

Qy        292 VLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQA 351
              ||||: |||||||||||||::|::|:|::||:| :|||||:| |||| :: :|||||
Db        130 VLMWLLTYVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQVDQYLGLVRTHINTVVAKIQA 189

Qy        352 KIPGLK 357
              |||| :
Db        190 KIPGAR 195
```

RESULT 4
T26216
hypothetical protein W06A7.3c - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26216
R;Ainscough, R.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z20173
A;Accession: T26216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2484 <WIL>
A;Cross-references: EMBL:Z78066; PIDN:CAB51467.1; GSPDB:GN00023; CESP:W06A7.3c
A;Experimental source: clone W06A7
C;Genetics:
A;Gene: CESP:W06A7.3c
A;Map position: 5
A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2

  Query Match 24.6%; Score 454.3; DB 2; Length 2484;
  Best Local Similarity 15.7%; Pred. No. 3.6e-06;
  Matches 126; Conservative 79; Mismatches 142; Indels 457; Gaps 18;



RESULT 5
T26215
hypothetical protein W06A7.3a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26215
R;Ainscough, R.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z20173
A;Accession: T26215
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2607 <WIL>
A;Cross-references: EMBL:Z78066; PIDN:CAB01522.2; GSPDB:GN00023; CESP:W06A7.3a
A;Experimental source: clone W06A7
C;Genetics:
A;Gene: CESP:W06A7.3a
A;Map position: 5
A;Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2;
2586/2

Query Match 24.2%; Score 447; DB 2; Length 2607;
Best Local Similarity 14.0%; Pred. No. 7.3e-06;
Matches 128; Conservative 79; Mismatches 150; Indels 560; Gaps 18;

```
Qy          1 MEDLDQSPLVSSSDSPPR-------------PQPAFKYQFV---------REPEDEEEEE 38
              :| |::   : :   ||               |    :|: |           |::::||
Db       1679 VESLERPLTIITQQKPPEKPTEDIGALSPLSPNTLAEYEEVPMMDMQSVPHSPQEKQEEI 1738

Qy         39 E-------------------------------------------------------EEE 42
              |                                                        |:
Db       1739 EALSEIIEEPQAMKEVEKPVESAPEKDNESLEAPEIINEPIRRVLVETKIMGPGKSLNED 1798

Qy         43 EDEDED---------------------------------------------------- 48
               |:|:|
Db       1799 NDDDDDGSECLDSIGDLSERTIQRFNTSIDDPSIRRDSFSSISSFGDRQKFRTAIENIRQ 1858

Qy         49 ----------------------------------------------LEELEVLER 57
                                                            ||:|:  :|
Db       1859 DLLPFQSSVSQYLRSSPNPSQQLLVTNLSMDSPSDLSPNAPPVGFENTAQFLEKLQQEDR 1918

Qy         58 KPAAG----------------LSAAPV------------------------------ 68
                | |                :| ||
Db       1919 PSAEGSIDSSGFEKVDHEGLDEFAAPPVHDPMQKSVFGSLGSDDMKPGSQDDGFVFIERN 1978

Qy         69 ----------------------------PTAPAAGAPLM------------------ 79
                                          |||    :|:
Db       1979 EANEATLKKNQKMSSHHNDVIEKNYFNDNAPTAALLESPIAEEARKLVQDAVESASEYKK 2038

Qy         80 ---------------------------------------DFGNDFVPPA---------- 89
                                                     || :: || |
Db       2039 QAVDSGDEIGRELLDNVEQKIEQVKEPIVDSLHKAYDGVGDFVHETVPNAVDDFVREAEK 2098

Qy         90 --PRGPLP-----------------------------AAPP-----VAP--ERQP---- 106
                |  |:|                            ||     |||  : :|
Db       2099 QLPESPVPEKIETPEPLVDIHDTVDKVHDEVDNFLRREPTPPFETDDVAPLSDDKPQFGN 2158
```

```
Qy        107 ----------------------------------SWDPSPVSSTVPAPSPLSAAAVSPSK 132
                                                :||   |:|  |      || :
Db       2159 QTPEEDETTFDRKGPLTIPEEVEKAAAAQNNDLDDFDPLVTSNTGAAFGAAVGAAAAVES 2218

Qy        133 LPEDD-------EPPARPPPPP---------PASV--------SPQAEP----------V 158
               | |::       |   ||| ||         |::|        |  : |
Db       2219 LTEEEMFGHQKFETVPRPPTPPKDISDEDVKPSTVNLGPSHHHSHPSSPHHSILKHHGDA 2278

Qy        159 W-----TPPAPAPAAPP------------------------------------------- 170
               |      ||    |  |
Db       2279 WIDFKTVPPCAQNAFSPGEIMFLLAFFVYLSCFASFFSKSLPLLDNLLSLVVYLSISLII 2338

Qy        171 ------------------------------------------------------------ 170

Db       2339 HVKHHRKFRWNEEQATTMSKLGAVGRGLYALIAFIVNIVLRVGLNVALVVGVAVSAHEAY 2398

Qy        171 ----------STSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVT 220
                       |:|::|||| ||: :|   :| :|  |  : :::|  |  |  |
Db       2399 KLTKSSGVLRKKEVLDVIYWRDAKKSAIVLSLALLVLFVLAKYPLLTVVTYSLLLALGAA 2458

Qy        221 ISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFL 280
               ||::| |   |:|:|  |||   |  :: : :| |   ::  : |  |   :|::|
Db       2459 AGFRVFKKVEAQIKKTDSEHPFSEILAQDLTLPQEKVHAQADVFVEHATCIANKLKKLVF 2518

Qy        281 VDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANK 340
               |:  ::|:|| :::|  ||: : |:| || || |: :|||| :|| :|  || :|   :
Db       2519 VESPLESIKFGLVLWSLTYIASWFSGFTLAILGLLGVFSVPKVYESNQEAIDPHLATISG 2578

Qy        341 NVKDAMAKIQAKIPGLK 357
               ::|:    |  |:| |:
Db       2579 HLKNVQNIIDEKLPFLR 2595
```

RESULT 6
T26213
hypothetical protein W06A7.3b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26213
R;Ainscough, R.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z20173
A;Accession: T26213
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-222 <WIL>
A;Cross-references: EMBL:Z78066; PIDN:CAB01523.1; GSPDB:GN00023; CESP:W06A7.3b
A;Experimental source: clone W06A7
C;Genetics:
A;Gene: CESP:W06A7.3b
A;Map position: 5
A;Introns: 27/1; 77/2; 201/2

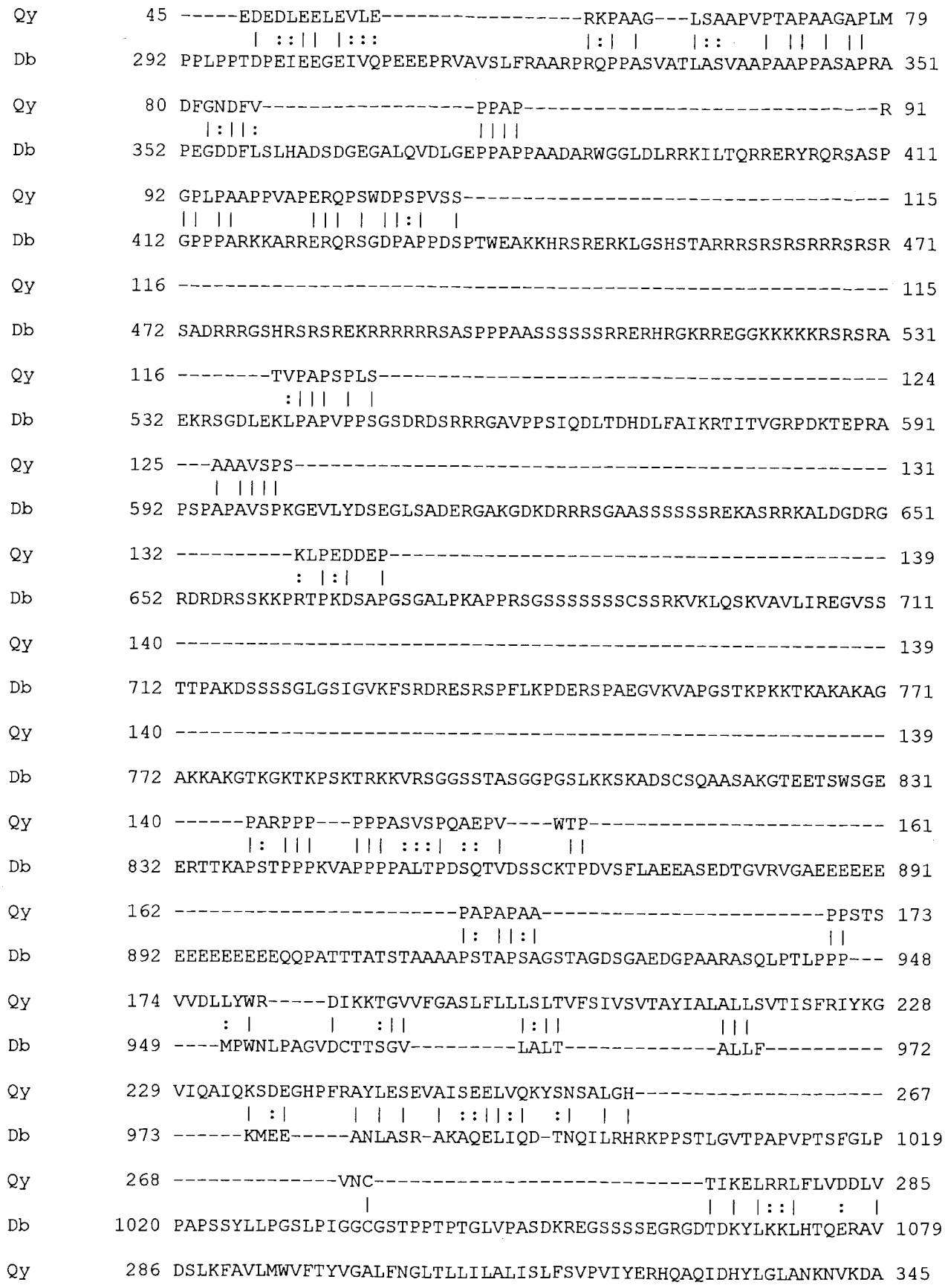
Query Match 19.1%; Score 353.5; DB 2; Length 222;
Best Local Similarity 32.2%; Pred. No. 2.2e-05;
Matches 65; Conservative 48; Mismatches 84; Indels 5; Gaps 1;

```
Qy        156 EPVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALA 215
              :| | |    |       ::|::|||| ||: :|   :| :|  |  : :::|  |  |
Db         14 QPTWVPATDFPG-----KILDVIYWRDAKKSAIVLSLALLVLFVLAKYPLLTVVTYSLLL 68

Qy        216 LLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKEL 275
               |     ||::| |   |:|:|  |||   |  :: : :| |   ::  : |  |   :|
Db         69 ALGAAAGFRVFKKVEAQIKKTDSEHPFSEILAQDLTLPQEKVHAQADVFVEHATCIANKL 128

Qy        276 RRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYL 335
              ::|  |:  ::|:|| :::|  ||: : |:| || || |: :|||| :|| :|  || :|
Db        129 KKLVFVESPLESIKFGLVLWSLTYIASWFSGFTLAILGLLGVFSVPKVYESNQEAIDPHL 188

Qy        336 GLANKNVKDAMAKIQAKIPGLK 357
                 : ::|:    |  |:| |:
Db        189 ATISGHLKNVQNIIDEKLPFLR 210
```

RESULT 7
T31421
C-terminal domain-binding protein rA1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999
C;Accession: T31421
R;Yuryev, A.; Patturajan, M.; Litingtung, Y.; Joshi, R.V.; Gentile, C.; Gebara, M.; Corden, J.L.
Proc. Natl. Acad. Sci. U.S.A. 93, 6975-6980, 1996
A;Title: The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins.
A;Reference number: Z21024; MUID:96293459; PMID:8692929
A;Accession: T31421
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1173 <YUR>
A;Cross-references: EMBL:U49056; NID:g1438531; PID:g1438532; PIDN:AAC52657.1
A;Experimental source: hippocampus

Query Match 14.9%; Score 275.1; DB 2; Length 1173;
Best Local Similarity 10.8%; Pred. No. 0.72;
Matches 130; Conservative 43; Mismatches 93; Indels 939; Gaps 30;

```
Qy          3 DLDQSPLV---------------------------------------------------- 10
              | |  ||
Db         52 DRDPMPLPVPSLLPRLRAWRTGKTVSPQSHASRPACSRHLLTLGTGDGGPAPPPAPSSGS 111

Qy         11 --------SSSDSPPRPQP----------------------AF--------KYQ----- 26
                     ||| ||| | |                      |:        ||
Db        112 SSPSPSPSSSSPSPPPPPPPPPPALPAPRFDIYDPFHPTDEAYSPPPAPEQKYDPFEAT 171

Qy         27 --------FVREPEDEEEEEEEEED---------------------------------- 44
                        ||:||||||||||:
Db        172 GSNPSSSGGTPSPEEEEEEEEEEEEGLSQSIRRISETLAGIYDDNSLSQDFPGDDSPHR 231

Qy         45 ------------------------------------------------------------ 44

Db        232 EPPPPQTLGAPGTPPQADSTRAEGAPRRRVFVVGPEAEACLEGKVSVEVVTTAGGPALPL 291
```



```
              : :|                                |:   |::     : |        ||
Db       1080 EEVK---------------------------LSIKPYYQKKDITKEEY--------KDI 1103

Qy        346 MAKIQAKI---------------------------------------------------- 353
              : |   ||
Db       1104 LRKAVHKICHSKSGEINPVKVSNLVRAYVQRYRYFRKHGRKPGDPPGPPRPPKEPGPPDK 1163

Qy        354 --PGL 356
                |||
Db       1164 GGPGL 1168
```

RESULT 8
S24407
formin isoform IV - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: S24407
R;Jackson-Grusby, L.; Kuo, A.; Leder, P.
Genes Dev. 6, 29-37, 1992
A;Title: A variant limb deformity transcript expressed in the embryonic mouse limb defines a novel formin.
A;Reference number: S24407; MUID:92112033; PMID:1339380
A;Accession: S24407
A;Molecule type: mRNA
A;Residues: 1-1206 <JAC>
A;Cross-references: EMBL:X62379; NID:g51552; PIDN:CAA44244.1; PID:g51553

  Query Match   14.4%; Score 266; DB 2; Length 1206;
  Best Local Similarity   17.3%; Pred. No. 1.6;
  Matches 102; Conservative   40; Mismatches   97; Indels 350; Gaps   24;

```
Qy          5 DQSPLVSSSDSPPR-----PQPAFKYQFVREPEDEEE------------------EEEEE 41
              ||||      :| |     | | |         | ||                    : | :
Db        459 DQSPTEQDDRTPGRLQAVWPPPKTK--------DTEEKVGLKYTEAEYQAAILHLKREHK 510

Qy         42 EEDE------------------------DEDLEEL-EVLERK----------------- 58
              || |                         :| :| | : ||::
Db        511 EEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKRREGCEEMRDVCISTDDDC 570

Qy         59 -PAA-------------------------------------------------------- 61
               | |
Db        571 SPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISLTQLSPSKDSKDIHAPFQT 630

Qy         62 --GLSA---------APVPTAPAAGAPLMDFGNDFVPPAP-----RGPLPAAPPVAPERQ 105
                | |:         || || |    ||       :|| |      |||| |||: |
Db        631 REGTSSSSQQKISPPAP-PTPPPLPPPL-------IPPPPPLPPGLGPLPPAPPIPP--- 679

Qy        106 PSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPA--------SVSPQAEP 157
                  |||   | | |       |: :|  | ||  ||||||         |  |   |
Db        680 ----VCPVSPPPPPPPP------PPTPVPPSDGPPPPPPPPPPLPNVLALPNSGGPPPPP 729

Qy        158 VWTPPAP--APAAPPSTSV--------------------VDLLYWRDIKKTGVVFGASLF 195
                 || |  ||  ||  |                     :  |||
Db        730 PPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKPLYWTR------------- 776

Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHP----------FRAY 245
```

```
                                                ||   || :  |           |
Db        777 ----------------------------------IQINDKSQDAAPTLWDSLEEPHIRDT 802

Qy        246 LESEVAISEELVQ------------------------KYSNSA---LGHVNCTIKELRR 277
                | |   |::  |                        | | :    :  ::  :|::::
Db        803 SEFEYLFSKDTTQQKKKPLSEAYEKKNKVKKIIKLLDGKRSQTVGILISSLHLEMKDIQQ 862

Qy        278 -LFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG 336
               :| |||                              ::: | ::  :||
Db        863 AIFTVDD----------------------------SVVDLETLAALYE---------- 882

Qy        337 LANKNVKDAMAKIQ------------------------AKIPGLKRKAE 361
                |:  :| : ||:                        |:||    :|:
Db        883 --NRAQEDELTKIRKYYETSKEEDLKLLDKPEQFLHELAQIPNFAERAQ 929


RESULT 9
S11515
formin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: S11515
R;Woychik, R.P.; Maas, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.
Nature 346, 850-853, 1990
A;Title: 'Formins': proteins deduced from the alternative transcripts of the
limb deformity gene.
A;Reference number: S11515; MUID:90363291; PMID:2392150
A;Accession: S11515
A;Molecule type: mRNA
A;Residues: 1-1468 <WOY>
A;Cross-references: EMBL:X53599; NID:g52877; PIDN:CAA37668.1; PID:g52878

  Query Match             14.4%;  Score 266;  DB 2;  Length 1468;
  Best Local Similarity   17.3%;  Pred. No. 2.6;
  Matches  102;  Conservative   40;  Mismatches   97;  Indels  350;  Gaps   24;

Qy          5 DQSPLVSSSDSPPR-----PQPAFKYQFVREPEDEEE-----------------EEEEE 41
              ||||      :| |     | |  |        | ||                  : | :
Db        685 DQSPTEQDDRTPGRLQAVWPPPKTK--------DTEEKVGLKYTEAEYQAAILHLKREHK 736

Qy         42 EEDE-----------------------DEDLEEL-EVLERK----------------- 58
              || |                        :| :| | : ||::
Db        737 EEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKRREGCEEMRDVCISTDDDC 796

Qy         59 -PAA------------------------------------------------------- 61
               | |
Db        797 SPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISLTQLSPSKDSKDIHAPFQT 856

Qy         62 --GLSA---------APVPTAPAAGAPLMDFGNDFVPPAP-----RGPLPAAPPVAPERQ 105
                | |:         || || |    ||       :|| |      |||| |||: |
Db        857 REGTSSSSQQKISPPAP-PTPPPLPPPL-------IPPPPPLPPGLGPLPPAPPIPP--- 905

Qy        106 PSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPA--------SVSPQAEP 157
                   |||   | | |       |: :|  | ||  ||||||          |  |   |
Db        906 ----VCPVSPPPPPPPP------PPTPVPPSDGPPPPPPPPPPLPNVLALPNSGGPPPPP 955

Qy        158 VWTPPAP--APAAPPSTSV--------------------VDLLYWRDIKKTGVVFGASLF 195
```

```
                 || |  ||  ||  |                      :  |||
Db        956 PPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKPLYWTR------------ 1002

Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHP----------FRAY 245
                                           ||   || :  |           |
Db       1003 ----------------------------------IQINDKSQDAAPTLWDSLEEPHIRDT 1028

Qy        246 LESEVAISEELVQ-------------------------KYSNSA---LGHVNCTIKELRR 277
                | |   |::  |                       | | :    :  ::  :|::::
Db       1029 SEFEYLFSKDTTQQKKKPLSEAYEKKNKVKKIIKLLDGKRSQTVGILISSLHLEMKDIQQ 1088

Qy        278 -LFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG 336
                :| |||                              ::: | ::  :||
Db       1089 AIFTVDD----------------------------SVVDLETLAALYE----------- 1108

Qy        337 LANKNVKDAMAKIQ-----------------------AKIPGLKRKAE 361
                |:  :| : ||:                       |:||    :|:
Db       1109 --NRAQEDELTKIRKYYETSKEEDLKLLDKPEQFLHELAQIPNFAERAQ 1155


RESULT 10
T13286
cappuccino gene protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C;Accession: T13286
R;Emmons, S.; Phan, H.; Calley, J.; Chen, W.; James, B.; Manseau, L.
Genes Dev. 9, 2482-2494, 1995
A;Title: Cappuccino, a Drosophila maternal effect rene required for polarity of
the egg and embryo, is related to the vertebrate limb deformity locus.
A;Reference number: Z17651; MUID:96033799; PMID:7590229
A;Accession: T13286
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1058 <EMM>
A;Cross-references: EMBL:U34258; NID:g1061333; PID:g1061334; PIDN:AAC46925.1
C;Genetics:
A;Gene: capu
A;Cross-references: FlyBase:FBgn0000256

  Query Match            14.3%;  Score 264.1;  DB 2;  Length 1058;
  Best Local Similarity  16.1%;  Pred. No. 1.3;
  Matches  118;  Conservative   47;  Mismatches  117;  Indels  449;  Gaps   27;

Qy          9 LVSSSDSPPR--------------PQPAFKYQF--------VREPEDEEEEEEEEE---- 42
              |:|| | |||                |  |::        |:   | | : | |:
Db        398 LLSSVDEPPRRPSKRCVNLTELLNASEATVYEYNKTGAEGCVKSFTDAETQTESEDCEGT 457

Qy         43 -------------EDEDEDLEELEVLERKPAAGLSAAPVPTAPAAGAPLMDFGNDFV-PP 88
                          :   || |       || |    || |  |    |       || ||
Db        458 CKCGQSSTKVSDNKSAKEDGE-------KPHA---VAPPPPPPPPPLPA------FVAPP 501

Qy         89 APRGPLPAAPPVAPERQPSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPAR------ 142
                |  | |  ||:|    |   | |     | | | | :|  |        |||
Db        502 PPPPPPPPPPPLANYGAP---PPP-----PPPPPGSGSAPPP-------PPPAPIEGGGG 546

Qy        143 -PPPPPPASVSPQAEPV-----------------------------------WT------ 160
```

```
                 ||||||  |  ||    :                                        ||
Db        547 IPPPPPPMSASPSKTTISPAPLPDPAEGNWFHRTNTMRKSAVNPPKPMRPLYWTRIVTSA 606

Qy        161 PPAP-------------------------------APAAPPSTSVVDLLYWRDIKKT---- 186
              ||||                              | |||:|  :    | :|::|
Db        607 PPAPRPPSVANSTDSTENSGSSPDEPPAANGADAPPTAPPATKEI----WTEIEETPLDN 662

Qy        187 ----------------------------------------GVVF---------------- 190
                                                     |:::
Db        663 IDEFTELFSRQAIAPVSKPKELKVKRAKSIKVLDPERSRNVGIIWRSLHVPSSEIEHAIY 722

Qy        191 --------------------------------GASLFL------LLSLTVFSI----VSV 208
                                             |  : |       || ::: |:    :|
Db        723 HIDTSVVSLEALQHMSNIQATEDELQRIKEAAGGDIPLDHPEQFLLDISLISMASERISC 782

Qy        209 TAYIALALLSVTISFRIYKGVIQAIQ---------------------------------- 234
                : |    |||: ||  : | |  |
Db        783 IVFQAEFEESVTLLFRKLETVSQLSQQLIESEDLKLVFSIILTLGNYMNGGNRQRGQADG 842

Qy        235 ------------KSDEGH------------------------------------------ 240
                         || | |
Db        843 FNLDILGKLKDVKSKESHTTLLHFIVRTYIAQRRKEGVHPLEIRLPIPEPADVERAAQMD 902

Qy        241 -----------------------------------PFRAYLESEVAISEELVQKYSNSAL 265
                                                ||:: :|  |  ::: : |   |
Db        903 FEEVQQQIFDLNKKFLGCKRTTAKVLAASRPEIMEPFKSKMEEFVEGADKSMAKLHQS-- 960

Qy        266 GHVNCTIKELRRLFL--------------------------------VDDLVDSLKFAVL 293
                    : | | |||                                 :|  |  |
Db        961 ------LDECRDLFLETMRFYHFSPKACTLTLAQCTPDQFFEYWTNFTNDFKDIWK---- 1010

Qy        294 MWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK---IQ 350
                    : :| | |               : :  ||||:     : :||   : |   |
Db       1011 ----KEITSLLNEL---------------MKKSKQAQIE-----SRRNVSTKVEKSGRIS 1046

Qy        351 AKIPGLKRKAE 361
               |   | |:::
Db       1047 LKERMLMRRSK 1057
```

RESULT 11
T38236
hypothetical protein SPAC23A1.17 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38236
R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21780
A;Accession: T38236
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1611 <MUR>
A;Cross-references: EMBL:AL021813; PIDN:CAA16991.1; GSPDB:GN00066;
SPDB:SPAC23A1.17
A;Experimental source: strain 972h-; cosmid c23A1
C;Genetics:

A;Gene: SPDB:SPAC23A1.17
A;Map position: 1

Query Match 14.2%; Score 262.4; DB 2; Length 1611;
Best Local Similarity 9.8%; Pred. No. 4.4;
Matches 127; Conservative 57; Mismatches 112; Indels 1006; Gaps 29;

```
Qy          2 EDL--------DQSPLV----------SSSDSPPRP---------------------- 19
              |:|        |  | |          :||  | :|
Db        299 EELSKSQRVAKDDDPFVVSNTANSDEPASSSKPAKPLTDLNRAFSQRLNLDPQKPGKSQG 358

Qy         20 ------------------QPAFKYQFVREPEDEEE----------EEEEEEEDEDEDL 49
                              |   ::   ||||::|           |||:|:| |:| :
Db        359 EISEQEEDEYDDAESDEMHSPYSTHEPESEPEDQDEPSEKDDENKDVEEEQEQEQEEEQI 418

Qy         50 EELE-----------------------------------VLERKPAAG---------- 62
              :  |                                    | | ||
Db        419 DPEEAKRIALRERMAKMSGGIGMHVFGLPGLAAPIGRKNTLRRTPAKSSEEAKSTTNDSS 478

Qy         63 ------------------------------------------------------------ 62

Db        479 PPKDSSSTSTQPTEQSNAQQAPSPKEEERPLPSEPSQNQPAEYRDTPDTPRNIMPLPGLM 538

Qy         63 ----------------------------------------------------------LSA 65
                                                                          ||
Db        539 SADQPIKVTEPSNDADKAIVAEGPNNEEETKGPVIPETQETSEQQVHKTPSPEKQKVLSP 598

Qy         66 APV--------------------PTAP------------------------------- 72
               |:                    |:||
Db        599 PPIITNFDKETLASNEAHEAVPQKPSAPQVTRLMAPQDSSSVVTPSPTSLLDPARAVRKV 658

Qy         73 --------------------AAGAPL---------------------------------- 78
                                 || :|:
Db        659 IDGIDPPKEAGAGATADVESAANSPITPPRTWHSPDFTSKSFEPIERKLPSRISEVTEDS 718

Qy         79 ----------------------MDFGN------------DFVPPAPR------------- 91
                                   : ||              | :|  ||
Db        719 IDEDKQNEVDPSTSARALPPPGLRFGKVDTLASLAHDDLDDLPAVPRIFSPPPLPKTPSG 778

Qy         92 ----------------------------------------------------------GP 93
                                                                          |
Db        779 EFGDNEFMFPKKSNRVRGHQSRPSTGSQLRNVVPVSIVTSGGRPALPDEMASPSSSIGHP 838

Qy         94 LPAAPPV------------------APERQPSWD-----------PSPVSST------- 116
              ||: ||                  |||  |||::          |:| ::|
Db        839 LPSPPPADFNSLNVDFYEPHSYLESPAPEPQPSYEEESFNATVIHAPTPSTATFQGHPTI 898

Qy        117 ------------------------------------------------------------ 116

Db        899 SNVATPPLKQDVTESKASPVADASATHQSSTGLTQEITQLGSNMRLPTKLTRPSNDGRKA 958

Qy        117 ----------VPAPSPLSAAAVSPSKLPEDDEPPARP---------------------- 143
                      :| | |:|    ||:  |  | ||: |
Db        959 SGPRPAAPPSIPPPLPVSNILSSPTSEPPKDHPPSAPLSKPVSTSPAAPLARVPPVPKLS 1018

Qy        144 ---------------------------------------PPPPPASVSPQAE--- 156
                                                      ||| ||   | :|
```

```
Db       1019 SKAPPVPLPSADAPPIPVPSTAPPVPIPTSTPPVPKSSSGAPSAPPPVPA---PSSEIPS 1075

Qy        157 ---PVWTPPAPAP---------------------------------------------- 166
                 |   || |||
Db       1076 IPAPSGAPPVPAPSGIPPVPKPSVAAPPVPKPSVAVPPVPAPSGAPPVPKPSVAAPPVPV 1135

Qy        167 --------------------------------AAPP------------------------ 170
                                               ||||
Db       1136 PSGAPPVPKPSVAAPPVPAPSGAPPVPKPSVAAPPVPAPSSGIPPVPKPAAGVPPVPPPS 1195

Qy        171 ------------------------------------------------------------ 170

Db       1196 EAPPVPKPSVGVPPVPPPSTAPPVPTPSAGLPPVPVPTAKAPPVPAPSSEAPSVSTPRSS 1255

Qy        171 ------------------------------------------------------------ 170

Db       1256 VPSPHSNASPSPTSSSMASAAPARTSVSRSKSKAERHETSTSSRKSSKSGEHHHHHNEGH 1315

Qy        171 --STSVVDLLYWRDIKKTGVVFGASLFLLLSLTVF------SIVSVTAYI---------- 212
                 |:|    |  :| :|       ||   || :        |||| |
Db       1316 ADSSSTRTSLAHQDSRK-------SLHRHLSRSSSRASKKPSIVSTTGPFNESFSAKPVE 1368

Qy        213 ----------------ALALLSVTISFRIYKGV--------------------------- 229
                             ::  :: :: : | :|:
Db       1369 PCASEKWWLNSTAVPKSVVQMNDSVLYMIKEGITGQDKKYKSVHILFPDYSQTVLTATFN 1428

Qy        230 --------IQAIQKSDEGHPFRAYLESEVAISEELVQK----YSNSALGHVNCTIKELRR 277
                     :  :| :    | :| |: | |     : |    |  | :|  :
Db       1429 PHNQNITQLSQLQLAPPAQPSKARLDEEYACYGSTILKKARAYQGSMVGDGS-------- 1480

Qy        278 LFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYL-- 335
                               ||:| :: :     |||                   | |
Db       1481 -----------------AFTFVNSVMS-----ILA-----------------HNLEP 1498

Qy        336 -------GLANKNV-------------KDAMAKIQAKIPGLK 357
                    |:  |||              | :   :||  | |
Db       1499 INKQTFGGVIYKNVGNVTVQQIGEIRPGDIVTFDKAKFSGQK 1540
```

RESULT 12
F86387
probable Pto kinase interactor [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F86387
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Altaf, H.; Araujo, R.; Bowman, C.L.; Brooks, S.Y.; Buehler, E.; Chan, A.; Chao, Q.; Chen, H.; Cheuk, R.F.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Dunn, P.; Etgu, P.; Feldblyum, T.V.; Feng, J.; Fong, B.; Fujii, C.Y.; Gill, J.E.; Goldsmith, A.D.; Haas, B.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.J.; Koo, H.L.; Kremenetskaia, I.; Kurtz, D.B.; Kwan, A.; Lam, B.; Langin-Hooper, S.; Lee, A.; Lee, J.M.; Lenz, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, A.; Militscher, J.; Miranda,

M.; Nguyen, M.; Nierman, W.C.; Osborne, B.I.; Pai, G.; Peterson, J.; Pham, P.K.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, L.J.; Tambunga, G.; Toriumi, M.J.; Town, C.D.; Utterback, T.; van Aken, S.; Vaysberg, M.; Vysotskaia, V.S.; Walker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86387
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-760 <STO>
A;Cross-references: GB:AE005172; NID:g11079512; PIDN:AAG29223.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 14.1%; Score 260.9; DB 2; Length 760;
Best Local Similarity 15.2%; Pred. No. 0.72;
Matches 119; Conservative 57; Mismatches 127; Indels 481; Gaps 27;

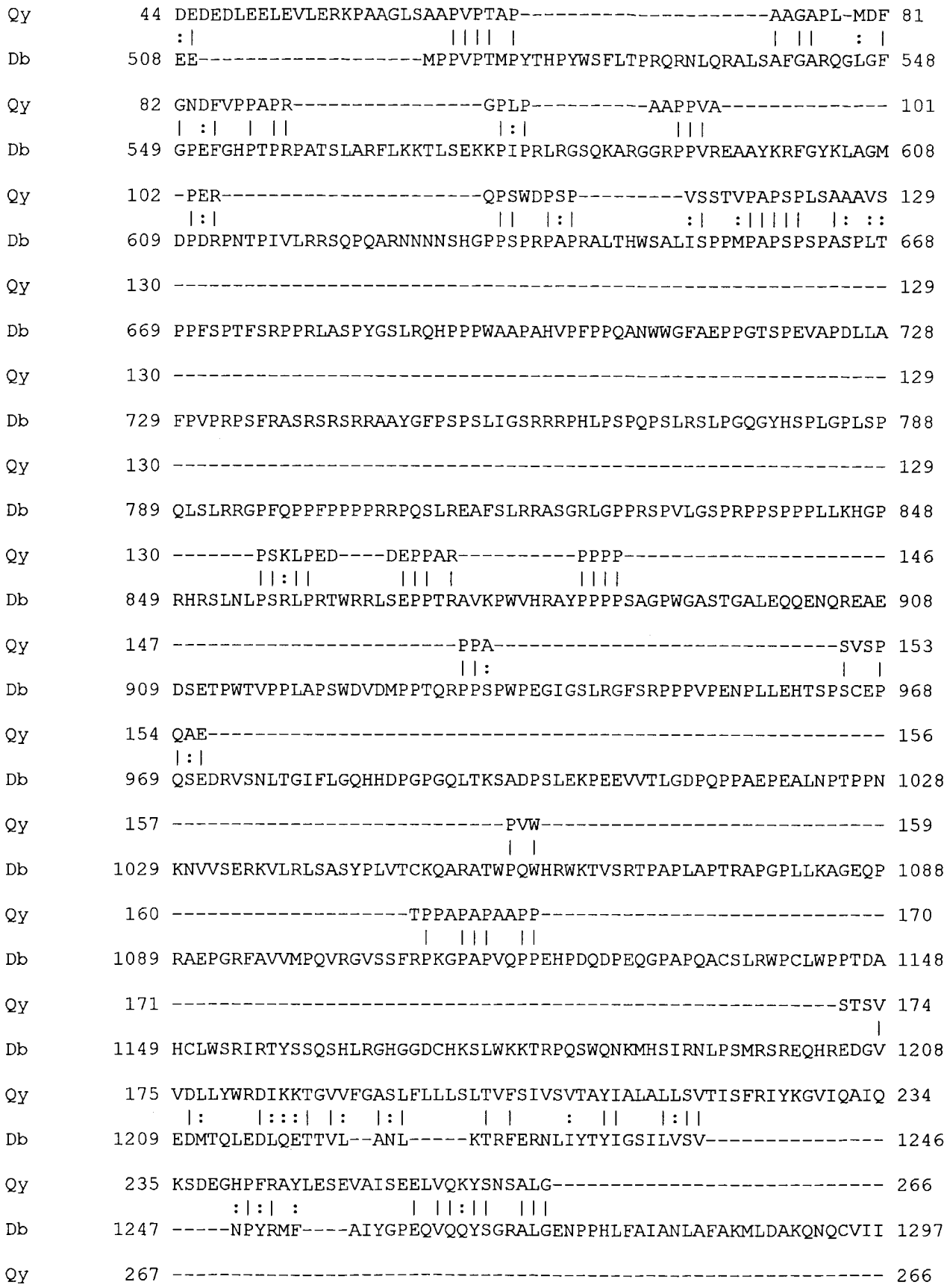
```
Qy          2 EDLDQSPLVSSSDSP-----PRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLE 56
              |::  ||   |  ||     | |||:|       |
Db         10 EEVSLSP---SLASPPLMALPPPQPSFPGDNATSP----------------------T 42

Qy         57 RKPAAG---------LSAAPVPTAPAAGAPLMDFGNDFVPPAP--RGPLPAAPPVAPERQ 105
              |:|  |           ::| |  | :  |      :  ||:|   || |   ||:|  :
Db         43 REPTNGNPPETTNTPAQSSPPPETPLSSPP-----PEPSPPSPSLTGPPPTTIPVSPPPE 97

Qy        106 PSWDP----------SPVSS---------------------------------------TV 117
              ||  |          :||||                                       ::
Db         98 PSPPPPLPTEAPPPANPVSSPPPESSPPPPPPTEAPPTTPITSPSPPTNPPPPPESPPSL 157

Qy        118 PAPSPLS-------------------------------------AAAVSPSKLPEDDEPP 140
              |||  | |                                    |:  ||  | | | |
Db        158 PAPDPPSNPLPPPKLVPPSHSPPRHLPSPPASEIPPPPRHLPSPPASERPSTPPSDSEHP 217

Qy        141 ARPP---------PPPPASVSPQAEP-------------------VWTPPAPAP----- 166
              : ||          |||| |  |   |                    || |:|
Db        218 SPPPPGHPKRREQPPPPGSKRPTPSPPSPSDSKRPVHPSPPSPPEETLPPPKPSPDPLPS 277

Qy        167 --------------AAPPS----------------TSVVDLLYWRDIKKTGVVFGASLF 195
                          :|||                 | | |     :   :|:
Db        278 NSSSPPTLLPPSSVVSPPSPPRKSVSGPDNPSPNNPTPVTD-----NSSSSGI------- 325

Qy        196 LLLSLTVFSIVSVTAYIALALLSV---------------------------------- 219
                  :: ::| |:  :|| ||::
Db        326 -----SIAAVVGVSIGVALVLLTLIGVVVCCLKKRKKRLSTIGGGYVMPTPMESSSPRSD 380

Qy        220 ------------------------------------------TISF---------- 223
                                                       |  |
Db        381 SALLKTQSSAPLVGNRSSNRTYLSQSEPGGFGQSRELFSYEELVIATNGFSDENLLGEGG 440

Qy        224 --RIYKGV-----IQAIQKSDEG-----HPFRAYLESEVAISEELVQKYSNSALGHVNCT 271
                |:||||     : |:::   |       |:| :::       : : :  : |  |
Db        441 FGRVYKGVLPDERVVAVKQLKIGGGQGDREFKAEVDT-------ISRVHHRNLLSMVGYC 493

Qy        272 IKELRRLFLVD----------------DLVDSLKFAV---------------------- 292
```

```
              | | ||| : |                  |    :| |
Db        494 ISENRRLLIYDYVPNNNLYFHLHGTPGLDWATRVKIAAGAARGLAYLHEDCHPRIIHRDI 553

Qy        293 ---------------------------------LMWVFTYVG---ALFNGLT------- 308
                                                :|  | |:    |    ||
Db        554 KSSNILLENNFHALVSDFGLAKLALDCNTHITTRVMGTFGYMAPEYASSGKLTEKSDVFS 613

Qy        309 --LLILALIS---------------------LFSVPVIYERHQAQIDHYLG-------- 336
                :::| ||:                    | |     |   |  |  ||
Db        614 FGVVLLELITGRKPVDASQPLGDESLVEWARPLLSNATETEEFTALADPKLGRNYVGVEM 673

Qy        337 ------------------------------LANKNVKDAM----------AKIQAKIPGL 356
                                           || ::: : |          |:  |:|
Db        674 FRMIEAAAACIRHSATKRPRMSQIVRAFDSLAEEDLTNGMRLGESEIINSAQQSAEIRLF 733

Qy        357 KRKA 360
              :| |
Db        734 RRMA 737
```

RESULT 13
A59295
unconventional myosin-15 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C;Accession: A59295
R;Liang, Y.; Wang, A.; Belyantseva, I.A.; Anderson, D.W.; Probst, F.J.; Barber, T.D.; Miller, W.; Touchman, J.W.; Jin, L.; Sullivan, S.L.; Sellers, J.R.; Camper, S.A.; Lloyd, R.V.; Kachar, B.; Friedman, T.B.; Fridell, R.A.
Genomics 61, 243-258, 1999
A;Title: Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2.
A;Reference number: A59266; MUID:20021762; PMID:10552926
A;Accession: A59295
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-3511 <LIA>
A;Cross-references: GB:AF144095; NID:g6224684; PIDN:AAF05904.1; PID:g6224685
C;Genetics:
A;Gene: MGI:Myo15
A;Cross-references: MGI:1261811
A;Map position: 11:33.9
C;Superfamily: myosin motor domain homology
F;1209-1871/Domain: myosin motor domain homology <MMO>

Query Match 13.9%; Score 257.1; DB 2; Length 3511;
Best Local Similarity 9.5%; Pred. No. 48;
Matches 136; Conservative 46; Mismatches 113; Indels 1139; Gaps 35;

```
Qy         16 PPRP-----QPAFKYQFV--------------------REPEDEEEE------------ 37
                |  |     :||| | :|                      |||:  ||
Db        388 PGMPYVYPEEPAFMYPWVPPPIMSPHNPYAHPMDDIAELEEPEETGEERQSTSFRLPSAA 447

Qy         38 ------------------------------------------------------EEEEEE 43
                                                                       | ||||
Db        448 FFEQQGMDKPARSKLSLIRKFRLFPRPQVKLFGKEKLEVPLPPSLDIPLPLGDAEGEEEE 507
```

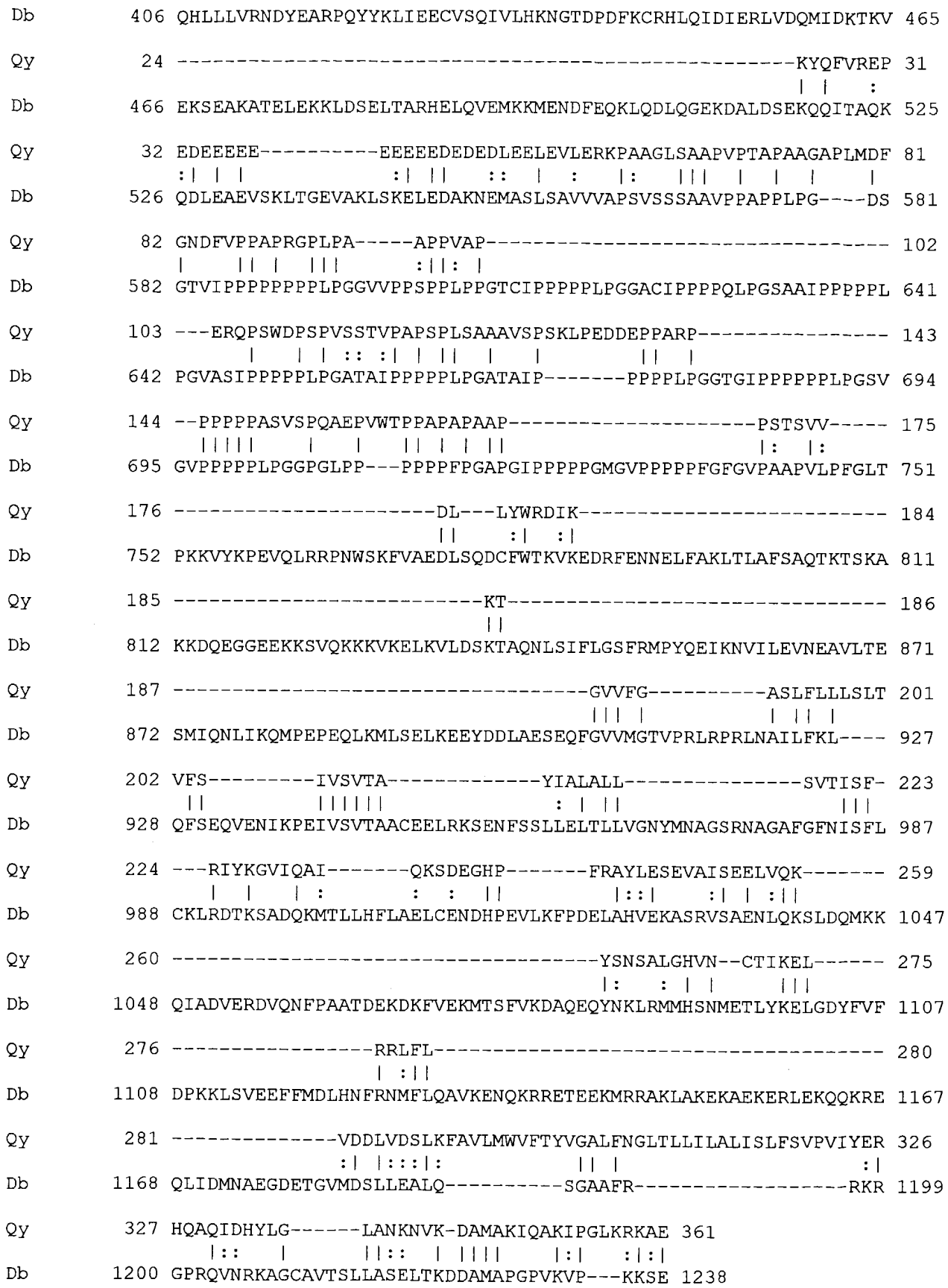


```
Db       1298 SGESGSGKTEATKLILRCLAAMNQRRDVMQQIKILEATPLLEAFGNAKTVRNDNSSRFGK 1357

Qy        267 ------------------------------------HV---------------------- 268
                                                  |:
Db       1358 FVEIFLEGGVICGAITSQYLLEKSRIVFQAKNERNYHIFYELLAGLPAQLRQAFSLQEAE 1417

Qy        269 ---------NCTI------KELRRLF---------------------------------- 279
                      || |       : |||
Db       1418 TYYYLNQGGNCEIAGKSDADDFRRLLAAMEVLGFTSEDQDSIFRILASILHLGNVYFEKH 1477

Qy        280 ------------------------------------------------------LVD--D 283
                                                                     ||  |
Db       1478 ETDAQEVASVVSAREIQAVAELLQVSPEGLQKAITFKVTETIREKIFTPLTVESAVDARD 1537

Qy        284 LVDSLKFAVLM-WVFTYVGAL----------------------FNGLTLLIL-----ALI 315
               :  : :|:|  |: | | ||                      ||    | :      |
Db       1538 AIAKVLYALLFGWLITRVNALVSPKQDTLSIAILDIYGFEDLSFNSFEQLCINYANENLQ 1597

Qy        316 SLFSVPVIYERHQAQ---------------------------------------------- 330
               ||:  ::::  | :
Db       1598 YLFN-KIVFQEEQEEYIREQMDWREIAFADNQPCINLISLKPYGILRILDDQCCFPQATD 1656

Qy        331 ----------------------------IDHYLGLANKNV-----------KDAM---- 346
                                          | || |     |           :| :
Db       1657 HTFLQKCHYHHGANPLYSKPKMPLPEFTIKHYAGKVTYQVHKFLDKNHDQVRQDVLDLFV 1716

Qy        347 ------------------------------------AKIQAKIPGLKRKAE 361
                                                  || |  :  |  | |
Db       1717 HSRTRVVAHLFSSHAAQTAPPRLGKSSSITRLYKAHTVAAKFQQSLLDLVEKME 1770
```

RESULT 14
T31065
diaphanous protein homolog p140mDia - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31065
R;Watanabe, N.; Madaule, P.; Reid, T.; Ishizaki, T.; Watanabe, G.; Kakizuka, A.; Saito, Y.; Nakao, K.; Jockusch, B.M.; Narumiya, S.
EMBO J. 16, 3044-3056, 1997
A;Title: P140mDia, a mammalian homolog of Drosophila diaphanous, is a target protein for Rho small GTPase and is a ligand for profilin.
A;Reference number: Z20961; MUID:97357293; PMID:9214622
A;Accession: T31065
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1255 <WAT>
A;Cross-references: EMBL:U96963; NID:g2114472; PID:g2114473; PIDN:AAC53280.1
A;Note: binds to GTP-bound form of Rho and binds to profilin

Query Match 13.9%; Score 256.5; DB 2; Length 1255;
Best Local Similarity 14.7%; Pred. No. 3.7;
Matches 130; Conservative 42; Mismatches 135; Indels 575; Gaps 31;

```
Qy          6 QSPLVSSSDSPPRPQ--------------------PAF---------------------- 23
              |  |:  :|   |||                    | |
```



RESULT 15
A59266
unconventional myosin-15 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 08-Sep-2000
C;Accession: A59266
R;Liang, Y.; Wang, A.; Belyantseva, I.A.; Anderson, D.W.; Probst, F.J.; Barber, T.D.; Miller, W.; Touchman, J.W.; Jin, L.; Sullivan, S.L.; Sellers, J.R.; Camper, S.A.; Lloyd, R.V.; Kachar, B.; Friedman, T.B.; Fridell, R.A.
Genomics 61, 243-258, 1999
A;Title: Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2.
A;Reference number: A59266; MUID:20021762; PMID:10552926
A;Accession: A59266
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-3530 <LIA>
A;Cross-references: GB:AF144094; NID:g6224682; PIDN:AAF05903.1; PID:g6224683
C;Superfamily: myosin motor domain homology
F;1225-1887/Domain: myosin motor domain homology <MMO>

Query Match 13.8%; Score 255.5; DB 2; Length 3530;
Best Local Similarity 9.9%; Pred. No. 55;
Matches 140; Conservative 41; Mismatches 125; Indels 1105; Gaps 36;

```
Qy         16 PPRPQPAFKYQF-------VREPEDEEEEE-------------EEE-------------- 41
                || | |   |         : ||||   |               |::
Db        416 PPIPSPHNPYAHAMDDIAELEEPEDAGVERQGTSFRLPSAAFFEQQGMDKPARSKLSLIR 475

Qy         42 ---------------------------------EEDEDEDLEEL----------------- 52
                                                : ||:|| |||
Db        476 KFRLFPRPQVKLFGKEKLEVPLPPSLDIPLPLGDADEEEDEEELPPVSAVPYGHPFWGFL 535

Qy         53 ---------------------------------------------EVLERKPAAGLS----- 64
                                                           : |:|| | |
Db        536 TPRQRNLQRALSAFGAHRGLGFGPEFGRPVPRPATSLARFLKKTLSEKKPIARLRGSQKT 595

Qy         65 ------------------------------------------------------------ 64

Db        596 RAGGPAVREAAYKRFGYKLAGMDPEKPGTPIVLRRAQPRARSSNDARRPPAPQPAPRTLS 655

Qy         65 ---------------------------------------------------AAP--VPTAPA 73
                                                                  |||  || ||
Db        656 HWSALLSPPVPPRPPSSGPPPAPPLSPALSGLPRPASPYGSLRRHPPPWAAPAHVPPAPQ 715

Qy         74 AG-------------------------------------------------------APLMD 80
              |                                                        :||
Db        716 ASGWAFVEPPAVSPEVPPDLLAFPGPRPSFRGSRRRGAAFGFPGASPRASRRRAWSPLAS 775

Qy         81 -----------------------------FGNDFVPPA--------------------- 89
                                           |   |:|||
Db        776 PQPSLRSSPGLGYCSPLAPPSPQLSLRTGPFQPPFLPPARRPRSLQESPAPRRAAGRLGP 835

Qy         90 -----PRGPLPAAPPV------------------------------------------- 100
                  |  | | :||:
```

```
Db        836 PGSPLPGSPRPPSPPLGLCHSPRRSSLNLPSRLPHTWRRLSEPPTRAVKPQVRLPFHRPP 895

Qy        101 --------------------------APERQPSWD---------------------------- 109

Db        896 RAGAWRAPLEHRESPREPEDSETPWTVPPLAPSWDVDMPPTQRPPSPWPGGAGSRRGFSR 955

Qy        110 PSPVSSTV------PAPSP----------------------------LSAAAVSPSK---- 132

Db        956 PPPVPENPFLQLLGPVPSPTLQPEDPAADMTRVFLGRHHEPGPGQLTKSAGPTPEKPEEE 1015

Qy        133 -------LPEDDEPPARPPP---------------------------------------- 145

Db       1016 ATLGDPQLPAETKPPTPAPPKDVTPPKDITPPKDVLPEQKTLRPSLSYPLAACDQTRATW 1075

Qy        146 PP-------------------------------------------------------PASV 151

Db       1076 PPWHRWGTLPQAAAPLAPIRAPEPLPKGGERRQAAPGRFAVVMPRVQKLSSFQRVGPATL 1135

Qy        152 SPQAEPVWTP--------------------PAP-----------APAA------------ 168

Db       1136 KPQVQPIQDPKPRACSLRWSCLWLRADAYGPWPRVHTHPQSCHLGPGAACLSLRGSWEEV 1195

Qy        169 -PPS----------------------TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSI 205

Db       1196 GPPSWRNKMHSIRNLPSMRFREQHGEDGVEDMTQLEDLQET--------------TVLSN 1241

Qy        206 VSVT-------AYIALALLSVT--ISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEEL 256

Db       1242 LKIRFERNLIYTYIGSILVSVNPYQMFGIY-------------------------GPEQ 1275

Qy        257 VQKYSNSALG---------------------------------------------------- 266

Db       1276 VQQYNGRALGENPPHLFAVANLAFAKMLDAKQNQCIIISGESGSGKTEATKLILRYLAAM 1335

Qy        267 ---------------------------------------------------------- 266

Db       1336 NQKREVMQQIKILEATPLLESFGNAKTVRNDNSSRFGKFVEIFLEGGVISGAITSQYLLE 1395

Qy        267 --------------HV-------------------------------NCTI------KEL 275

Db       1396 KSRIVFQAKNERNYHIFYELLAGLPAQLRQAFSLQEAETYYYLNQGGNCEIAGKSDADDF 1455

Qy        276 RRLF------------------------------------------------------- 279

Db       1456 RRLLAAMEVLGFSSEDQDSIFRILASILHLGNVYFEKYETDAQEVASVVSAREIQAVAEL 1515

Qy        280 --------------------------------LVD--DLVDSLKFAVLM-WVFTYVGAL- 303

Db       1516 LQISPEGLQKAITFKVTETMREKIFTPLTVESAVDARDAIAKVLYALLFSWLITRVNALV 1575

Qy        304 ---------------------FNGLTLLIL-----ALISLFSVPVIYERHQ----AQID- 332

Db       1576 SPRQDTLSIAILDIYGFEDLSFNSFEQLCINYANENLQYLFNKIVFQEEQEEYIREQIDW 1635

Qy        333 ---------------------------------------------------------- 332

Db       1636 QEITFADNQPRINLISLKPYGILRILDDQCCFPQATDHTFLQKCHYHHGANPLYSKPKMP 1695
```

```
Qy        333 -------HYLGLANKNV------------KDAM--------------------------- 346
                     || |     |             :| :
Db       1696 LPEFTIKHYAGKVTYQVHKFLDKNHDQVRQDVLDLFVRSRTRVVAHLFSSHAPQAAPQRL 1755

Qy        347 ----------------AKIQAKIPGLKRKAE 361
                              || |  :  |  | |
Db       1756 GKSSSVTRLYKAHTVAAKFQQSLLDLVEKME 1786
```

Search completed: September 29, 2004, 18:55:49
Job time : 33.5368 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:54:54 ; Search time 86.1193 Seconds
(without alignments)
1348.937 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description

1	1833.8	99.1	373	9	US-09-789-386-6	Sequence 6, Appli
2	1833.8	99.1	373	9	US-09-765-205-6	Sequence 6, Appli
3	1833.8	99.1	373	9	US-09-893-348-24	Sequence 24, Appl
4	1833.8	99.1	373	12	US-10-408-967-8	Sequence 8, Appli
5	1833.8	99.1	373	14	US-10-060-036-72	Sequence 72, Appl
6	1825.8	98.7	373	16	US-10-466-258-4	Sequence 4, Appli
7	1756.9	95.0	1192	9	US-09-789-386-2	Sequence 2, Appli
8	1756.9	95.0	1192	9	US-09-758-140-6	Sequence 6, Appli
9	1756.9	95.0	1192	9	US-09-893-348-23	Sequence 23, Appl
10	1756.9	95.0	1192	9	US-09-972-599A-6	Sequence 6, Appli
11	1756.9	95.0	1192	12	US-10-267-502-429	Sequence 429, App
12	1756.9	95.0	1192	14	US-10-060-036-71	Sequence 71, Appl
13	1756.9	95.0	1192	16	US-10-327-213-9	Sequence 9, Appli
14	1756.9	95.0	1192	16	US-10-466-258-9	Sequence 9, Appli
15	1749.9	94.6	1192	12	US-10-408-967-7	Sequence 7, Appli
16	1556.2	84.1	379	14	US-10-205-194-164	Sequence 164, App
17	1546.3	83.6	360	9	US-09-893-348-20	Sequence 20, Appl
18	1478.8	79.9	1163	9	US-09-893-348-18	Sequence 18, Appl
19	1441.2	77.9	1163	12	US-10-267-502-431	Sequence 431, App
20	927	50.1	199	9	US-09-893-348-25	Sequence 25, Appl
21	927	50.1	199	12	US-10-660-946-1	Sequence 1, Appli
22	920	49.7	199	12	US-10-408-967-9	Sequence 9, Appli
23	918	49.6	199	12	US-09-978-360A-467	Sequence 467, App
24	904	48.9	199	9	US-09-893-348-21	Sequence 21, Appl
25	886	47.9	199	16	US-10-466-258-11	Sequence 11, Appl
26	827.7	44.7	776	12	US-10-660-946-5	Sequence 5, Appli
27	827.7	44.7	776	12	US-10-267-502-430	Sequence 430, App
28	825.8	44.6	780	12	US-10-267-502-432	Sequence 432, App
29	818.1	44.2	777	14	US-10-205-219-93	Sequence 93, Appl
30	789.8	42.7	356	12	US-10-660-946-6	Sequence 6, Appli
31	684.2	37.0	208	12	US-10-660-946-7	Sequence 7, Appli
32	682.6	36.9	593	15	US-10-108-260A-2892	Sequence 2892, Ap
33	666.8	36.0	266	12	US-10-276-774-2330	Sequence 2330, Ap
34	666	36.0	267	12	US-10-660-946-8	Sequence 8, Appli
35	666	36.0	267	14	US-10-205-194-127	Sequence 127, App
36	662.6	35.8	269	14	US-10-106-698-6222	Sequence 6222, Ap
37	644.2	34.8	236	9	US-09-729-674-20	Sequence 20, Appl
38	644.2	34.8	236	9	US-09-765-205-26	Sequence 26, Appl
39	644.2	34.8	236	12	US-10-408-967-2	Sequence 2, Appli
40	541.9	29.3	168	10	US-09-809-391-563	Sequence 563, App
41	541.9	29.3	168	10	US-09-882-171-563	Sequence 563, App
42	541.9	29.3	168	12	US-10-164-861-563	Sequence 563, App
43	538.3	29.1	241	12	US-10-660-946-3	Sequence 3, Appli
44	527.9	28.5	222	12	US-10-267-502-428	Sequence 428, App
45	517.4	28.0	234	12	US-10-424-599-200840	Sequence 200840,

ALIGNMENTS

RESULT 1
US-09-789-386-6
; Sequence 6, Application US/09789386
; Patent No. US20020010324A1
; GENERAL INFORMATION:
; APPLICANT: MICHALOVICH, DAVID

```
;  APPLICANT:  PRINJHA, RABINDER KUMAR
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP-30165-C1
;  CURRENT APPLICATION NUMBER: US/09/789,386
;  CURRENT FILING DATE:  2001-02-21
;  PRIOR APPLICATION NUMBER: U.K. 9916898.1
;  PRIOR FILING DATE: 1999-07-19
;  PRIOR APPLICATION NUMBER: U.K. 9816024.5
;  PRIOR FILING DATE: 1998-07-22
;  PRIOR APPLICATION NUMBER: US 09/359,208
;  PRIOR FILING DATE: 1999-07-22
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-789-386-6

  Query Match             99.1%;  Score 1833.8;  DB 9;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 8.6e-60;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373


RESULT 2
US-09-765-205-6
; Sequence 6, Application US/09765205
; Patent No. US20020034800A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Cao, Li
;  TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
;  FILE REFERENCE: 1458.004/200130.449
;  CURRENT APPLICATION NUMBER: US/09/765,205
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: US/09/212,440
;  PRIOR FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: human
US-09-765-205-6

  Query Match             99.1%;  Score 1833.8;  DB 9;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 8.6e-60;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373


RESULT 3
US-09-893-348-24
; Sequence 24, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
;  APPLICANT: EISENBACH-SCHWARTZ, Michal
;  APPLICANT:  COHEN, Irun R.
;  APPLICANT:  BESERMAN, Pierre
;  APPLICANT:  MOSONEGO, Alon
```

; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: IL 124500
; PRIOR FILING DATE: 1998-05-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-348-24

Query Match 99.1%; Score 1833.8; DB 9; Length 373;
Best Local Similarity 96.5%; Pred. No. 8.6e-60;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 4
US-10-408-967-8

; Sequence 8, Application US/10408967
; Publication No. US20040063161A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Yan, Riqiang
; APPLICANT: Lu, Yifeng
; TITLE OF INVENTION: Compositions and Methods of Treating Alzheimer's Disease
; FILE REFERENCE: 00925
; CURRENT APPLICATION NUMBER: US/10/408,967
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-967-8

Query Match 99.1%; Score 1833.8; DB 12; Length 373;
Best Local Similarity 96.5%; Pred. No. 8.6e-60;
Matches 360; Conservative 0; Mismatches 1; Indels 12; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373
```

RESULT 5
US-10-060-036-72
; Sequence 72, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.

```
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Persing, David H.
;  APPLICANT:  Hepler, William T.
;  APPLICANT:  Jiang, Yuqiu
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF PANCREATIC CANCER
;  FILE REFERENCE: 210121.566
;  CURRENT APPLICATION NUMBER: US/10/060,036
;  CURRENT FILING DATE:  2002-01-30
;  NUMBER OF SEQ ID NOS: 4560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 72
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-060-036-72

  Query Match             99.1%;  Score 1833.8;  DB 14;  Length 373;
  Best Local Similarity   96.5%;  Pred. No. 8.6e-60;
  Matches  360;  Conservative    0;  Mismatches    1;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373


RESULT 6
US-10-466-258-4
; Sequence 4, Application US/10466258
; Publication No. US20040132096A1
; GENERAL INFORMATION:
;  APPLICANT: GLAXO GROUP LIMITED
```

```
;  TITLE OF INVENTION: ASSAY
;  FILE REFERENCE: P80966 GCW
;  CURRENT APPLICATION NUMBER: US/10/466,258
;  CURRENT FILING DATE:  2003-07-15
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-466-258-4

  Query Match             98.7%;  Score 1825.8;  DB 16;  Length 373;
  Best Local Similarity   96.0%;  Pred. No. 1.7e-59;
  Matches  358;  Conservative    1;  Mismatches    2;  Indels   12;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPCWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------- 173
              ||||||||||||||:||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPQDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 240

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 300

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 360

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
Db        361 IQAKIPGLKRKAE 373


RESULT 7
US-09-789-386-2
; Sequence 2, Application US/09789386
; Patent No. US20020010324A1
; GENERAL INFORMATION:
;  APPLICANT: MICHALOVICH, DAVID
;  APPLICANT:  PRINJHA, RABINDER KUMAR
;  TITLE OF INVENTION: NOVEL COMPOUNDS
;  FILE REFERENCE: GP-30165-C1
;  CURRENT APPLICATION NUMBER: US/09/789,386
;  CURRENT FILING DATE:  2001-02-21
;  PRIOR APPLICATION NUMBER: U.K. 9916898.1
```

```
;  PRIOR FILING DATE: 1999-07-19
;  PRIOR APPLICATION NUMBER: U.K. 9816024.5
;  PRIOR FILING DATE: 1998-07-22
;  PRIOR APPLICATION NUMBER: US 09/359,208
;  PRIOR FILING DATE: 1999-07-22
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-789-386-2

  Query Match             95.0%;  Score 1756.9;  DB 9;  Length 1192;
  Best Local Similarity   30.3%;  Pred. No. 4.9e-56;
  Matches  361;  Conservative    0;  Mismatches    0;  Indels  831;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
```

```
Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 8
US-09-758-140-6
; Sequence 6, Application US/09758140
; Patent No. US20020012965A1
; GENERAL INFORMATION:
; APPLICANT: Strittmatter, Stephen M.
; TITLE OF INVENTION: No. US20020012965A1o Receptor-Mediated Blockade of Axonal Growth
; FILE REFERENCE: 44574-5073-US
; CURRENT APPLICATION NUMBER: US/09/758,140
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,707
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: US 60/207,366
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/236,378

; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-758-140-6

  Query Match   95.0%; Score 1756.9; DB 9; Length 1192;
  Best Local Similarity   30.3%; Pred. No. 4.9e-56;
  Matches 361; Conservative   0; Mismatches   0; Indels 831; Gaps   1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
```

```
Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 9
US-09-893-348-23
; Sequence 23, Application US/09893348
; Patent No. US20020072493A1
; GENERAL INFORMATION:
; APPLICANT: EISENBACH-SCHWARTZ, Michal
; APPLICANT: COHEN, Irun R.
; APPLICANT: BESERMAN, Pierre
; APPLICANT: MOSONEGO, Alon
; APPLICANT: MOALEM, Gila
; TITLE OF INVENTION: ACTIVATED T-CELLS, NERVOUS SYSTEM-SPECIFIC ANTIGENS AND THEIR USES
; FILE REFERENCE: EIS-SCHWARTZ=2A
; CURRENT APPLICATION NUMBER: US/09/893,348
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/314,161
; PRIOR FILING DATE: 1999-05-19
; PRIOR APPLICATION NUMBER: US 09/218,277
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: PCT/US98/14715

```
;  PRIOR FILING DATE: 1998-07-21
;  PRIOR APPLICATION NUMBER: IL 124500
;  PRIOR FILING DATE: 1998-05-19
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-893-348-23

  Query Match             95.0%;  Score 1756.9;  DB 9;  Length 1192;
  Best Local Similarity   30.3%;  Pred. No. 4.9e-56;
  Matches  361;  Conservative    0;  Mismatches    0;  Indels  831;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171
```

```
Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 -------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                          ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 10
US-09-972-599A-6
; Sequence 6, Application US/09972599A
; Patent No. US20020077295A1
; GENERAL INFORMATION:
; APPLICANT: STRITTMATTER, STEPHEN M.
; TITLE OF INVENTION: NOGO-RECEPTOR-MEDIATED BLOCKADE OF AXONAL GROWTH
; FILE REFERENCE: C077 CIP US
; CURRENT APPLICATION NUMBER: US/09/972,599A
; CURRENT FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: PCT/US01/01041
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 09/758,140
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/236,378
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/207,366
; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/175,707
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-972-599A-6

Query Match 95.0%; Score 1756.9; DB 9; Length 1192;
Best Local Similarity 30.3%; Pred. No. 4.9e-56;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
```

```
Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 11
US-10-267-502-429
; Sequence 429, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 429
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-267-502-429

Query Match 95.0%; Score 1756.9; DB 12; Length 1192;
Best Local Similarity 30.3%; Pred. No. 4.9e-56;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

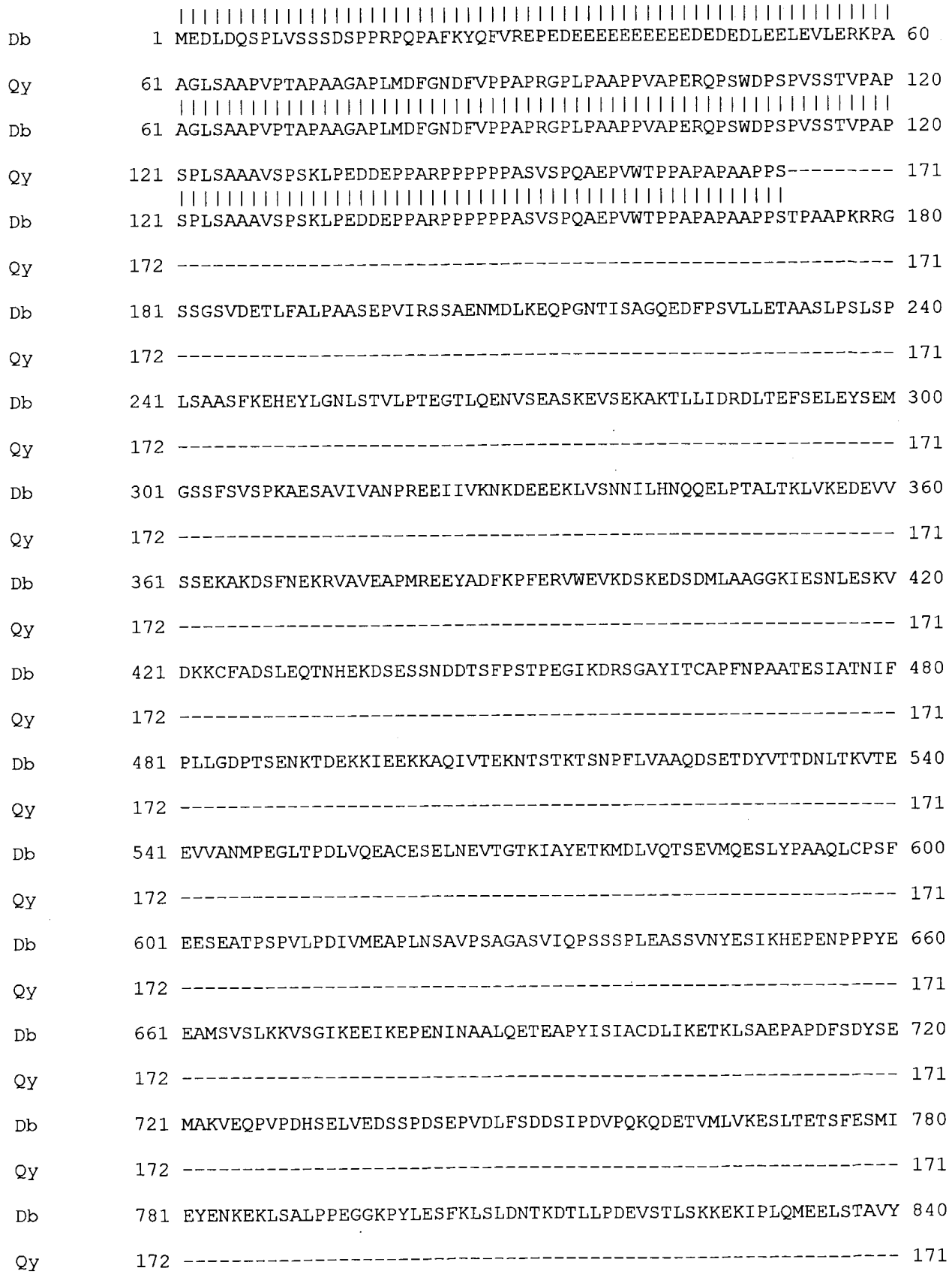
```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171
```

```
Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 12
US-10-060-036-71
; Sequence 71, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-71

Query Match 95.0%; Score 1756.9; DB 14; Length 1192;
Best Local Similarity 30.3%; Pred. No. 4.9e-56;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
```



```
Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 13
US-10-327-213-9
; Sequence 9, Application US/10327213
; Publication No. US20040121341A1
; GENERAL INFORMATION:
; APPLICANT: FILBIN, MARIE T.
; APPLICANT: DOMENICONI, MARCO
; APPLICANT: CAO, ZIXUAN
; TITLE OF INVENTION: INHIBITORS OF MYELIN-ASSOCIATED GLYCOPROTEIN (MAG)
; TITLE OF INVENTION: ACTIVITY FOR REGULATING NEURAL GROWTH AND REGENERATION
; FILE REFERENCE: CUNY/003
; CURRENT APPLICATION NUMBER: US/10/327,213
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-327-213-9

Query Match 95.0%; Score 1756.9; DB 16; Length 1192;
Best Local Similarity 30.3%; Pred. No. 4.9e-56;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171

Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ----------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
```

```
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192


RESULT 14
US-10-466-258-9
; Sequence 9, Application US/10466258
; Publication No. US20040132096A1
; GENERAL INFORMATION:
;  APPLICANT: GLAXO GROUP LIMITED
;  TITLE OF INVENTION: ASSAY
;  FILE REFERENCE: P80966 GCW
;  CURRENT APPLICATION NUMBER: US/10/466,258
;  CURRENT FILING DATE:  2003-07-15
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 1192
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-466-258-9

  Query Match             95.0%;  Score 1756.9;  DB 16;  Length 1192;
  Best Local Similarity   30.3%;  Pred. No. 4.9e-56;
  Matches  361;  Conservative    0;  Mismatches    0;  Indels  831;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300
```

```
Qy        172 ------------------------------------------------------------ 171
Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360
Qy        172 ------------------------------------------------------------ 171
Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420
Qy        172 ------------------------------------------------------------ 171
Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480
Qy        172 ------------------------------------------------------------ 171
Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540
Qy        172 ------------------------------------------------------------ 171
Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
Qy        172 ------------------------------------------------------------ 171
Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
Qy        172 ------------------------------------------------------------ 171
Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720
Qy        172 ------------------------------------------------------------ 171
Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780
Qy        172 ------------------------------------------------------------ 171
Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840
Qy        172 ------------------------------------------------------------ 171
Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900
Qy        172 ------------------------------------------------------------ 171
Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960
Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020
Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080
Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140
Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
```

```
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

RESULT 15
US-10-408-967-7
; Sequence 7, Application US/10408967
; Publication No. US20040063161A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn Company
; APPLICANT: Yan, Riqiang
; APPLICANT: Lu, Yifeng
; TITLE OF INVENTION: Compositions and Methods of Treating Alzheimer's Disease
; FILE REFERENCE: 00925
; CURRENT APPLICATION NUMBER: US/10/408,967
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-967-7

Query Match 94.6%; Score 1749.9; DB 12; Length 1192;
Best Local Similarity 30.2%; Pred. No. 8.9e-56;
Matches 360; Conservative 0; Mismatches 1; Indels 831; Gaps 1;

```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171
```

```
Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRPGAYITCAPFNPAATESIATNIF 480
Qy        172 ------------------------------------------------------------ 171
Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540
Qy        172 ------------------------------------------------------------ 171
Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600
Qy        172 ------------------------------------------------------------ 171
Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660
Qy        172 ------------------------------------------------------------ 171
Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720
Qy        172 ------------------------------------------------------------ 171
Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780
Qy        172 ------------------------------------------------------------ 171
Db        781 EYENKEKLSALPPEGGKPYLESFKLSLINTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840
Qy        172 ------------------------------------------------------------ 171
Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900
Qy        172 ------------------------------------------------------------ 171
Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960
Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                        ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020
Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFPAYLESE 1080
Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140
Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

Search completed: September 29, 2004, 19:11:36
Job time : 93.1193 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:40:14 ; Search time 76.6061 Seconds
(without alignments)
1486.854 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description

1	1833.9	99.1	392	4	Q96B16	Q96b16 homo sapien
2	1533	82.9	375	11	Q8BHF5	Q8bhf5 mus musculu
3	1523.1	82.3	356	11	Q8BH78	Q8bh78 mus musculu
4	1513	81.8	357	11	Q8K3G7	Q8k3g7 mus musculu
5	1455.3	78.7	1162	11	Q8BGM9	Q8bgm9 mus musculu
6	1441.2	77.9	1163	11	Q8K3G8	Q8k3g8 mus musculu
7	1028	55.6	986	4	Q8IUA4	Q8iua4 homo sapien
8	994	53.7	1046	11	Q8BGK7	Q8bgk7 mus musculu
9	989.2	53.5	639	11	Q8K290	Q8k290 mus musculu
10	986.8	53.3	578	11	Q80W95	Q80w95 mus musculu
11	901	48.7	184	6	Q7YRW9	Q7yrw9 bos taurus
12	883.2	47.7	199	13	Q7T224	Q7t224 gallus gall
13	867	46.9	179	6	Q9GM33	Q9gm33 macaca fasc
14	829.8	44.9	780	11	Q8K0T0	Q8k0t0 mus musculu
15	825.8	44.6	780	11	Q8K4S4	Q8k4s4 mus musculu
16	793.3	42.9	760	13	Q90638	Q90638 gallus gall
17	750.9	40.6	214	13	Q7T222	Q7t222 carassius a
18	705.1	38.1	643	11	Q8CCU2	Q8ccu2 mus musculu
19	683	36.9	199	4	Q9BQ59	Q9bq59 homo sapien
20	673.3	36.4	720	11	Q7TNB7	Q7tnb7 mus musculu
21	669.9	36.2	208	13	Q90637	Q90637 gallus gall
22	669.2	36.2	595	5	Q9VMV9	Q9vmv9 drosophila
23	666	36.0	267	11	Q63765	Q63765 rattus sp.
24	652.4	35.3	236	11	Q8VBU0	Q8vbu0 rattus norv
25	651.1	35.2	237	11	Q8C6D5	Q8c6d5 mus musculu
26	595.6	32.2	221	13	Q7ZUD6	Q7zud6 brachydanio
27	529.3	28.6	234	5	Q9VMW3	Q9vmw3 drosophila
28	527.9	28.5	222	5	Q9VMW4	Q9vmw4 drosophila
29	521	28.2	224	5	Q9VMW1	Q9vmw1 drosophila
30	518	28.0	202	5	Q9VMW2	Q9vmw2 drosophila
31	454.3	24.6	2484	5	Q9U347	Q9u347 caenorhabdi
32	447	24.2	2607	5	Q23187	Q23187 caenorhabdi
33	353.5	19.1	222	5	Q23188	Q23188 caenorhabdi
34	343	18.5	107	13	Q7T223	Q7t223 carassius a
35	286.7	15.5	1179	12	Q91L98	Q91l98 white spot
36	286.6	15.5	1180	12	Q8VAS9	Q8vas9 white spot
37	286.1	15.5	1185	12	Q8QTC5	Q8qtc5 white spot
38	277	15.0	1312	4	Q9NR59	Q9nr59 homo sapien
39	277	15.0	1312	4	Q7Z5V7	Q7z5v7 homo sapien
40	277	15.0	1343	4	Q9H7N4	Q9h7n4 homo sapien
41	277	15.0	1627	10	Q84ZL0	Q84zl0 oryza sativ
42	275.1	14.9	1173	11	Q63624	Q63624 rattus norv
43	267.3	14.4	2301	10	Q9ATK5	Q9atk5 chlamydomon
44	266.4	14.4	1049	5	Q8MRP5	Q8mrp5 drosophila
45	266.4	14.4	1154	5	Q8IQ12	Q8iq12 drosophila

ALIGNMENTS

RESULT 1
Q96B16
ID Q96B16 PRELIMINARY; PRT; 392 AA.
AC Q96B16;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```
DE   Hypothetical protein (RTN4).
GN   RTN4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22376540; PubMed=12488097;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   J. Mol. Biol. 325:299-323(2003).
DR   EMBL; BC016165; AAH16165.1; -.
DR   EMBL; AY102285; AAM64242.1; -.
DR   EMBL; AY102278; AAM64247.1; -.
DR   GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   392 AA;  42274 MW;  D7B2AA5E839E58AD CRC64;

  Query Match             99.1%;  Score 1833.9;  DB 4;  Length 392;
  Best Local Similarity   91.8%;  Pred. No. 3.4e-68;
  Matches  360;  Conservative    1;  Mismatches    0;  Indels   31;  Gaps    1;

Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPST-------- 172
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180
```

```
Qy        173 ----------------------SVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVT 209
                                    :|||||||||||||||||||||||||||||||||||||
Db        181 SSGSVDETLFALPAASEPVIRSSAVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVT 240

Qy        210 AYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVN 269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVN 300

Qy        270 CTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQA 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQA 360

Qy        330 QIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||
Db        361 QIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 392
```

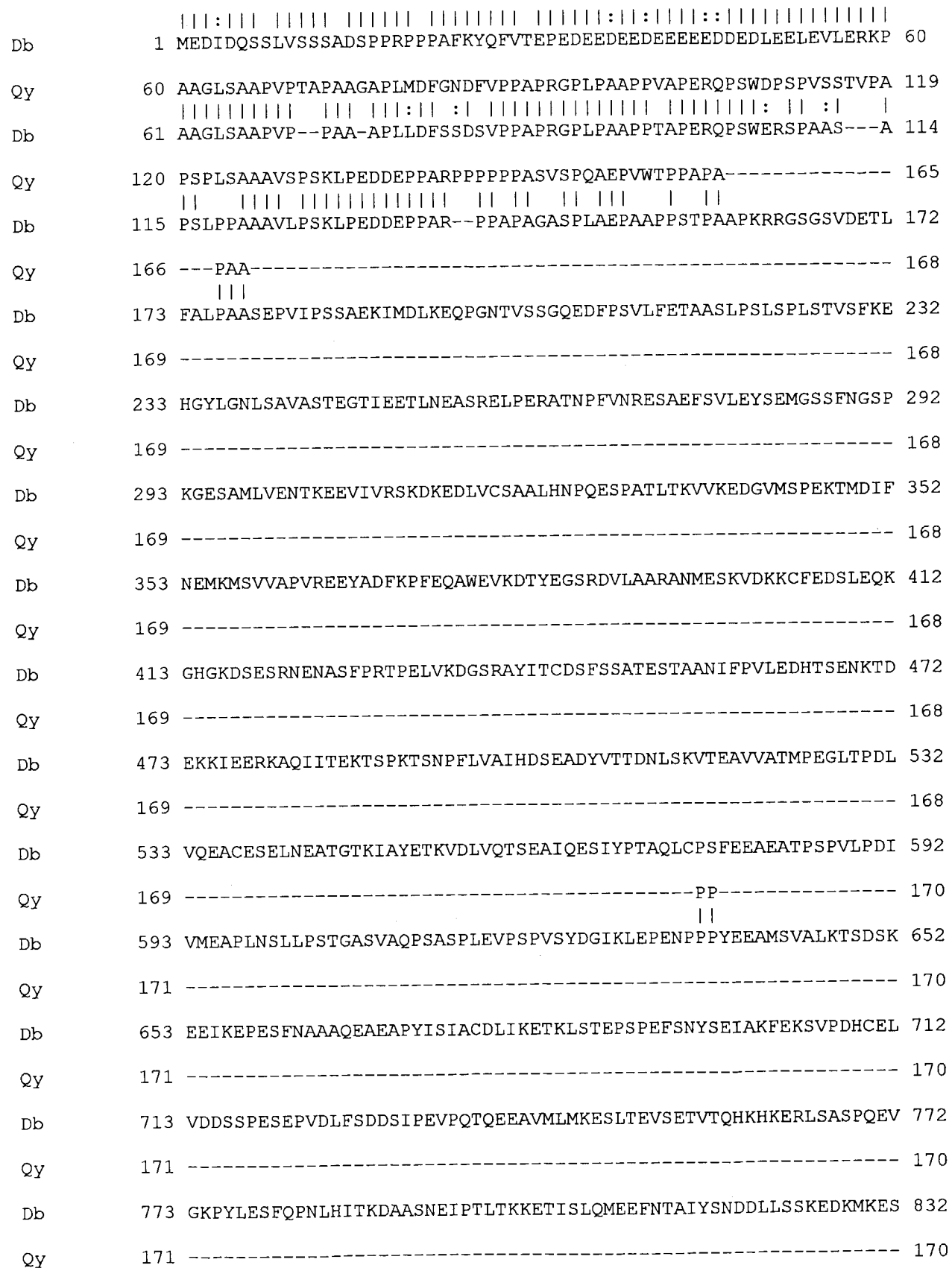
```
RESULT 2
Q8BHF5
ID   Q8BHF5      PRELIMINARY;      PRT;   375 AA.
AC   Q8BHF5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7;
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SvcJ7;
RA   Van der Putten H., Mir A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY102282; AAM73504.1; -.
DR   EMBL; AY102286; AAM73509.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
```

```
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   375 AA;  40300 MW;  23D9EB19BE671AE6 CRC64;

  Query Match             82.9%;  Score 1533;  DB 11;  Length 375;
  Best Local Similarity   82.8%;  Pred. No. 1.1e-55;
  Matches  317;  Conservative   13;  Mismatches   23;  Indels   30;  Gaps    7;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKP 59
              |||:||| ||||| |||||| |||||||| |||||:||:|||||::|||||||||||||||
Db          1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEEEEEDDEDLEELEVLERKP 60

Qy         60 AAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPA 119
              |||||||||| ||| |||:|| :| |||||||||||||| |||||||: || :|   |
Db         61 AAGLSAAPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAAS---A 114

Qy        120 PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPA-------------- 165
              ||   |||| ||||||||||||  || ||  || |||   |  ||
Db        115 PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAEPAAPPSTPAAPKRRGSGSVDETL 172

Qy        166 ---PAAP----PSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS 218
                 |||     ||::|||||||||||||||||||||||||||||||||||||||||||||
Db        173 FALPAASEPVIPSSAVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLS 232

Qy        219 VTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRL 278
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        233 VTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRL 292

Qy        279 FLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLA 338
              |||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        293 FLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLA 352

Qy        339 NKNVKDAMAKIQAKIPGLKRKAE 361
              ||:|||||||||||||||||||
Db        353 NKSVKDAMAKIQAKIPGLKRKAE 375


RESULT 3
Q8BH78
ID   Q8BH78      PRELIMINARY;      PRT;   356 AA.
AC   Q8BH78;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
```

```
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7;
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SvcJ7;
RA   Van der Putten H., Mir A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY102281; AAM73503.1; -.
DR   EMBL; AY102286; AAM73508.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   356 AA;  38403 MW;  4366C03BA9630B56 CRC64;

  Query Match             82.3%;  Score 1523.1;  DB 11;  Length 356;
  Best Local Similarity   84.5%;  Pred. No. 2.5e-55;
  Matches  315;  Conservative   11;  Mismatches   18;  Indels   29;  Gaps    7;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKP 59
              |||:||| ||||| |||||| |||||||| ||||||:||:||||::|||||||||||||
Db          1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEEEEEDDEDLEELEVLERKP 60

Qy         60 AAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPA 119
              |||||||||| ||| |||:|| :| |||||||||||||| |||||||: || :|   |
Db         61 AAGLSAAPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAAS---A 114

Qy        120 PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------ 173
              ||   |||| |||||||||||||  || ||  || ||         |||||||
Db        115 PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAE---------PAAPPSTPAAPKRR 163

Qy        174 -----VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 228
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GSGSVVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKG 223

Qy        229 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSL 288
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        224 VIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDSL 283

Qy        289 KFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK 348
              |||||||||||||||||||||||||||||:||||||||||||||||||:|||||||
Db        284 KFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAK 343

Qy        349 IQAKIPGLKRKAE 361
              |||||||||||||
```

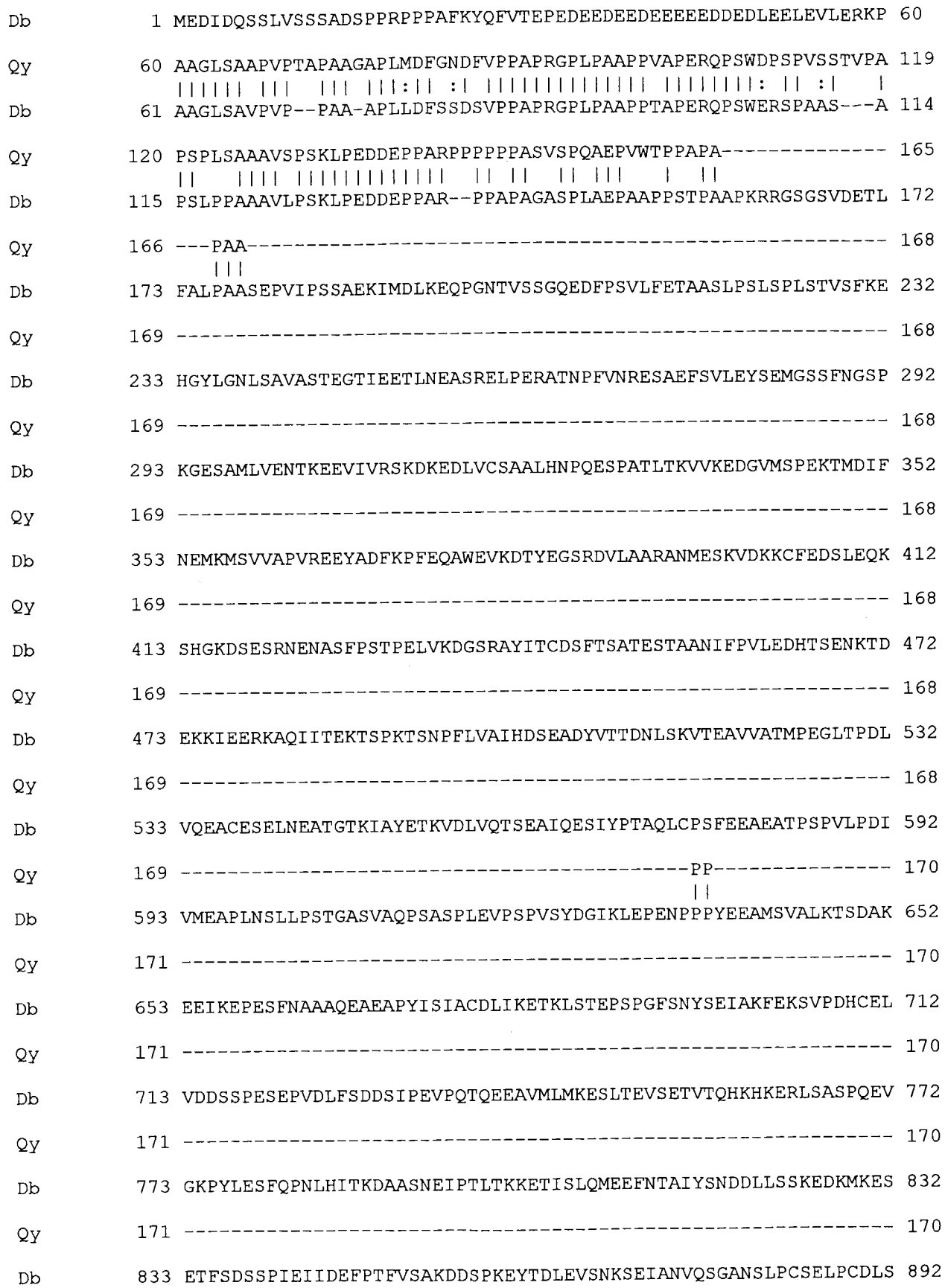
```
Db        344 IQAKIPGLKRKAE 356


RESULT 4
Q8K3G7
ID   Q8K3G7      PRELIMINARY;      PRT;   357 AA.
AC   Q8K3G7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nogo-B.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Jin W., Li R., Long M., Shen J., Ju G.;
RT   "Cloning and expression of the mouse Nogo-B protein.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY114153; AAM77069.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   357 AA;  38566 MW;  73BB3D17DFDBDF15 CRC64;

  Query Match           81.8%;  Score 1513;  DB 11;  Length 357;
  Best Local Similarity 84.2%;  Pred. No. 6.5e-55;
  Matches  315;  Conservative   11;  Mismatches   18;  Indels   30;  Gaps    8;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEVLERKP 59
              |||:||| ||||| |||||| |||||||| ||||||:||:|||||::|||||||||||||
Db          1 MEDIDQSSLVSSSADSPPRPPPAFKYQFVTEPEDEEDEEDEEEEEDDEDLEELEVLERKP 60

Qy         60 AAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPA 119
              |||||||||| ||| |||:|| :| |||||||||||||| ||||||||: || :|   |
Db         61 AAGLSAAPVP--PAA-APLLDFSSDSVPPAPRGPLPAAPPTAPERQPSWERSPAAS---A 114

Qy        120 PSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTS------ 173
              ||   |||| ||||||||||||  || ||  || ||         |||||||
Db        115 PSLPPAAAVLPSKLPEDDEPPAR--PPAPAGASPLAE---------PAAPPSTPAAPKRR 163

Qy        174 -----VVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYK 227
                  |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        164 GSGSVVVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYK 223

Qy        228 GVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDS 287
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        224 GVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDS 283

Qy        288 LKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMA 347
              |||||||||||||||||||||||||||||:||||||||||||||||||:||||||
```

```
Db        284 LKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMA 343

Qy        348 KIQAKIPGLKRKAE 361
              ||||||||||||||
Db        344 KIQAKIPGLKRKAE 357


RESULT 5
Q8BGM9
ID   Q8BGM9      PRELIMINARY;      PRT;  1162 AA.
AC   Q8BGM9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7;
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SvcJ7;
RA   Van der Putten H., Mir A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY102284; AAM73506.1; -.
DR   EMBL; AY102286; AAM73511.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   1162 AA;  126613 MW;  855697FBEE11781F CRC64;

  Query Match             78.7%;  Score 1455.3;  DB 11;  Length 1162;
  Best Local Similarity   27.3%;  Pred. No. 3.5e-51;
  Matches  319;  Conservative   11;  Mismatches   23;  Indels  817;  Gaps    8;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKP 59
```



```
Db        833 ETFSDSSPIEIIDEFPTFVSAKDDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLS 892

Qy        171 ------------------------------------------------------------ 170

Db        893 FKNTYPKDEAHVSDEFSKSRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEEKLP 952

Qy        171 -------------------STSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAY 211
                               |||||||||||||||||||||||||||||||||||||||||||
Db        953 SDTEKEDRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAY 1012

Qy        212 IALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCT 271
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1013 IALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNST 1072

Qy        272 IKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQI 331
              |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db       1073 IKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQI 1132

Qy        332 DHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||:||||||||||||||||||||
Db       1133 DHYLGLANKSVKDAMAKIQAKIPGLKRKAE 1162
```

```
RESULT 6
Q8K3G8
ID   Q8K3G8      PRELIMINARY;      PRT;  1163 AA.
AC   Q8K3G8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nogo-A.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Jin W., Long M., Li R., Ju G.;
RT   "Cloning and expression of the mouse Nogo-A protein.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY114152; AAM77068.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   1163 AA;  126691 MW;  6B5F362799417EA4 CRC64;

  Query Match           77.9%;  Score 1441.2;  DB 11;  Length 1163;
  Best Local Similarity 27.2%;  Pred. No. 1.4e-50;
  Matches  318;  Conservative   11;  Mismatches   24;  Indels  818;  Gaps    9;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKP 59
              |||:||| ||||| |||||| |||||||| |||||||:||:||||::|||||||||||||||
```

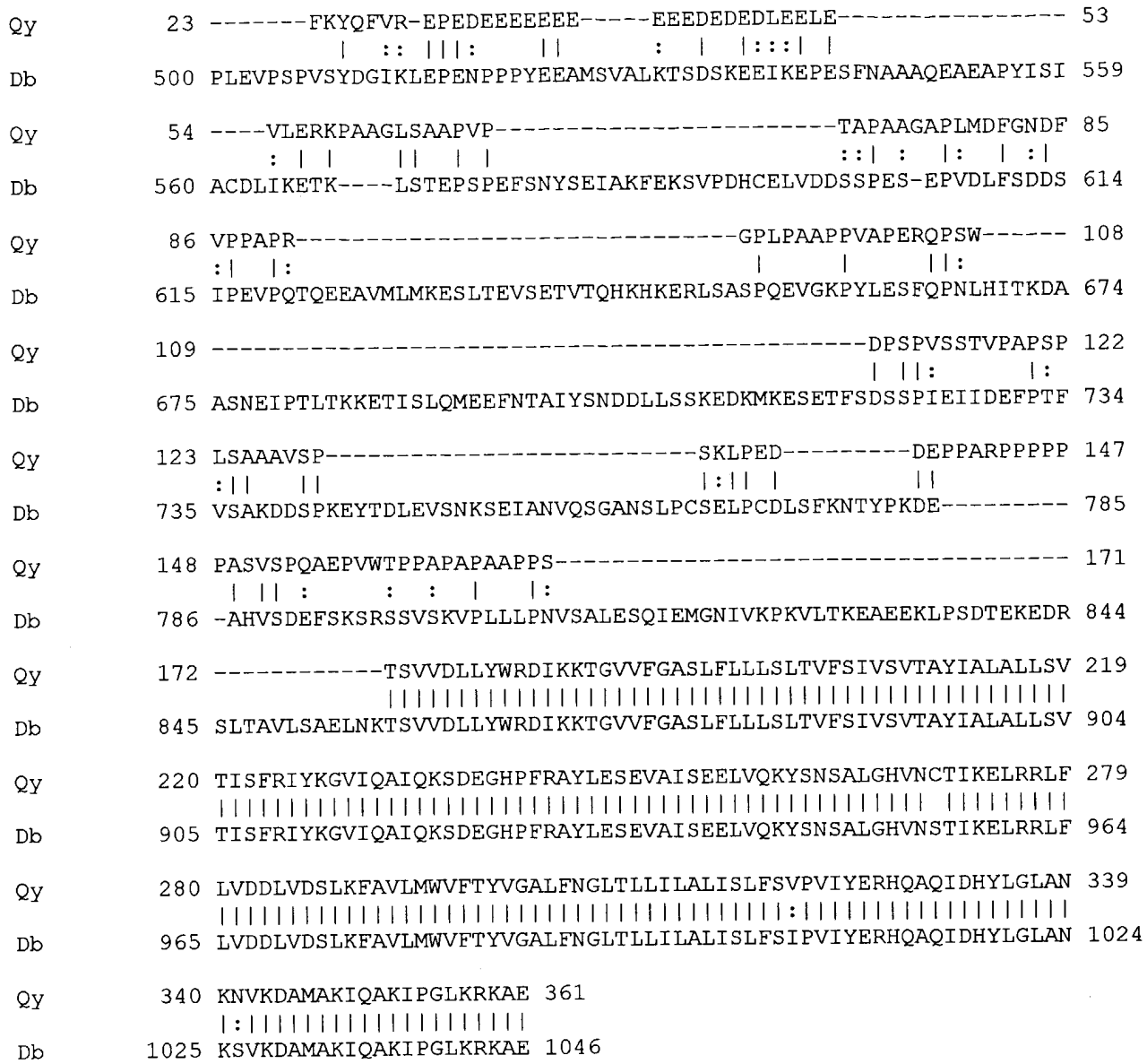


```
Qy        171 ------------------------------------------------------------ 170

Db        893 FKNTYPKDEAHVSDEFSKSRSSVSKVPLLLPNVSALESQIEMGNIVKPKVLTKEAEEKLP 952

Qy        171 -------------------STSVVDLLYWRDIKKTGVV-FGASLFLLLSLTVFSIVSVTA 210
                               ||||||||||||||||||| ||||||||||||||||||||
Db        953 SDTEKEDRSLTAVLSAELNKTSVVDLLYWRDIKKTGVVYFGASLFLLLSLTVFSIVSVTA 1012

Qy        211 YIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNC 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 YIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNS 1072

Qy        271 TIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQ 330
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1073 TIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQ 1132

Qy        331 IDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              ||||||||||:||||||||||||||||||||
Db       1133 IDHYLGLANKSVKDAMAKIQAKIPGLKRKAE 1163
```

```
RESULT 7
Q8IUA4
ID   Q8IUA4      PRELIMINARY;      PRT;   986 AA.
AC   Q8IUA4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   RNT4 (RTN4).
GN   RTN4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=22376540; PubMed=12488097;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   J. Mol. Biol. 325:299-323(2003).
```

```
DR   EMBL; AY102285; AAM64244.1; -.
DR   EMBL; AY123245; AAM64249.1; -.
DR   EMBL; AY123246; AAM64250.1; -.
DR   EMBL; AY123247; AAM64251.1; -.
DR   EMBL; AY123248; AAM64252.1; -.
DR   EMBL; AY123249; AAM64253.1; -.
DR   EMBL; AY123250; AAM64254.1; -.
DR   GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   986 AA;  108449 MW;  0CDE8F647036415A CRC64;

  Query Match             55.6%;  Score 1028;  DB 4;  Length 986;
  Best Local Similarity   39.2%;  Pred. No. 1.5e-33;
  Matches  250;  Conservative   27;  Mismatches   60;  Indels  300;  Gaps   17;

Qy          3 DLDQS---------------PLVSSSDSPPRP----------------------QP--- 21
              || |:               |    |:: | |                        ||
Db        372 DLVQTSEVMQESLYPAAQLCPSFEESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSS 431

Qy         22 -----AFKYQFVR-EPED----EE-----------EEEEEEEEDEDEDLEELE------ 53
                   :  |: :: |||:    ||            :|| :| |: :  |:| |
Db        432 PLEASSVNYESIKHEPENPPPYEEAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISI 491

Qy         54 ----VLERKPAAGLSAAPVP----------------------TAPAAGAPLMDFGNDF 85
                 : | |    ||| | |                       ::| :  |:  | :|
Db        492 ACDLIKETK----LSAEPAPDFSDYSEMAKVEQPVPDHSELVEDSSPDS-EPVDLFSDDS 546

Qy         86 VPPAP-------------------------RGPLPAAPP------------------ 99
              :|  |                        :  | | ||
Db        547 IPDVPQKQDETVMLVKESLTETSFESMIEYENKEKLSALPPEGGKPYLESFKLSLDNTKD 606

Qy        100 -VAP-----------------------------------ERQPSWDPSPV------- 113
               : |                                  | :   | ||:
Db        607 TLLPDEVSTLSKKEKIPLQMEELSTAVYSNDDLFISKEAQIRETETFSDSSPIEIIDEFP 666

Qy        114 ------------------------------------------------------------ 113

Db        667 TLISSKTDSFSKLAREYTDLEVSHKSEIANAPDGAGSLPCTELPHDLSLKNIQPKVEEKI 726

Qy        114 -------------SSTVPAPSPLSAAA--------VSP--------SKLPEDDEPPARPP 144
                          |  :  |  :|| |        | |         ||| | |   |
Db        727 SFSDDFSKNGSATSKVLLLPPDVSALATQAEIESIVKPKVLVKEAEKKLPSDTEKEDR-- 784

Qy        145 PPPPASVSPQAEPVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS 204
                          :| |   |    ||||||||||||||||||||||||||||||||
Db        785 ---------------SPSAIFSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS 829

Qy        205 IVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        830 IVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSA 889

Qy        265 LGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIY 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        890 LGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIY 949
```

```
Qy        325 ERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||
Db        950 ERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 986


RESULT 8
Q8BGK7
ID   Q8BGK7      PRELIMINARY;      PRT;  1046 AA.
AC   Q8BGK7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4.
GN   RTN4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., van der Putten H., Schwab M.E.;
RT   "Genomic Structure and Functional Characterization of the Promoter
RT   Structures of Human and Mouse Nogo/Rtn-4.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7, and 129SvcJ7;
RA   Oertle T., Schwab M.E.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvcJ7;
RA   Van der Putten H.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SvcJ7;
RA   Van der Putten H., Mir A.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY102280; AAM73502.1; -.
DR   EMBL; AY102286; AAM73507.1; -.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   1046 AA;  114221 MW;  8CE2E2238ED51222 CRC64;

  Query Match           53.7%;  Score 994;  DB 11;  Length 1046;
  Best Local Similarity 38.6%;  Pred. No. 4.7e-32;
  Matches  240;  Conservative   31;  Mismatches   73;  Indels  278;  Gaps   15;

Qy          3 DLDQS---------------PLVSSSDSPPRP----------------------QPA-- 22
              || |:               |    ::: | |                        ||:
Db        440 DLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSTGASVAQPSAS 499
```



RESULT 9
Q8K290
ID Q8K290 PRELIMINARY; PRT; 639 AA.
AC Q8K290;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN RTN4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC032192; AAH32192.1; -.
DR MGD; MGI:1915835; Rtn4.
DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR GO; GO:0007399; P:neurogenesis; IDA.
DR InterPro; IPR003388; Reticulon.
DR Pfam; PF02453; Reticulon; 1.
DR PROSITE; PS50845; RETICULON; 1.
KW Hypothetical protein.
SQ SEQUENCE 639 AA; 70312 MW; 309A19DA37603F11 CRC64;

Query Match 53.5%; Score 989.2; DB 11; Length 639;
Best Local Similarity 38.6%; Pred. No. 2.1e-32;
Matches 240; Conservative 31; Mismatches 73; Indels 278; Gaps 15;

```
Qy          3 DLDQS---------------PLVSSSDSPPRP----------------------QPA-- 22
              ||  |:                |    ::: | |                        ||:
Db         33 DLVQTSEAIQESIYPTAQLCPSFEEAEATPSPVLPDIVMEAPLNSLLPSTGASVAQPSAS 92

Qy         23 -------FKYQFVR-EPEDEEEEEEE-----EEEDEDEDLEELE--------------- 53
                        |  :: |||:    ||      :  |  |:::| |
Db         93 PLEVPSPVSYDGIKLEPENPPPYEEAMSVALKTSDAKEEIKEPESFNAAAQEAEAPYISI 152

Qy         54 ----VLERKPAAGLSAAPVP----------------------TAPAAGAPLMDFGNDF 85
                  : | |    ||  | |                      ::| :  |:  | :|
Db        153 ACDLIKETK----LSTEPSPEFSNYSEIAKFEKSVPDHCELVDDSSPES-EPVDLFSDDS 207

Qy         86 VPPAPR-------------------------------GPLPAAPPVAPERQPSW------ 108
              :|  |:                                |     |     ||:
Db        208 IPEVPQTQEEAVMLMKESLTEVSETVTQHKHKERLSASPQEVGKPYLESFQPNLHITKDA 267

Qy        109 ---------------------------------------------DPSPVSSTVPAPSP 122
                                                          | ||:      |:
Db        268 ASNEIPTLTKKETISLQMEEFNTAIYSNDDLLSSKEDKMKESETFSDSSPIEIIDEFPTF 327

Qy        123 LSAAAVSP--------------------------SKLPED---------DEPPARPPPPP 147
              :||   ||                        |:|| |          ||
Db        328 VSAKDDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNTYPKDE--------- 378

Qy        148 PASVSPQAEPVWTPPAPAPAAPPS---------------------------------- 171
               | || :     :  :  |   |:
Db        379 -AHVSDEFSKSRSSVSKVPLLLLPNVSALESQIEMGNIVKPKVLTKEAEEKLPSDTEKEDR 437

Qy        172 ------------TSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSV 219
                         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 SLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSV 497

Qy        220 TISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLF 279
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        498 TISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLF 557

Qy        280 LVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLAN 339
              |||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        558 LVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLAN 617

Qy        340 KNVKDAMAKIQAKIPGLKRKAE 361
```

```
                |:||||||||||||||||||||
Db        618 KSVKDAMAKIQAKIPGLKRKAE 639


RESULT 10
Q80W95
ID   Q80W95      PRELIMINARY;      PRT;   578 AA.
AC   Q80W95;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nogo-A (Fragment).
GN   NOGO-A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tozaki H., Hirata T.;
RT   "The partial sequence of mouse nogo-A cDNA clone#4109.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB073672; BAC75974.1; -.
DR   GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   578 AA;  63696 MW;  832670C171E4AC61 CRC64;

  Query Match           53.3%;  Score 986.8;  DB 11;  Length 578;
  Best Local Similarity 41.1%;  Pred. No. 2e-32;
  Matches  232;  Conservative   21;  Mismatches   85;  Indels  227;  Gaps    8;

Qy          7 SPLVSSSD--------SPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDLEELEV---- 54
              |||   |          |  | |  :   |     : :|| :| |  :   :| |
Db         31 SPLEVPSPVSYDGIKLEPENPPPYEEAMSVALKTSDAKEEIKEPESFNAAAQEAEAPYIS 90

Qy         55 ----------LERKPAAGLS----------AAP-----VPTAPAAGAPLMDFGNDFVPPA 89
                       |  :|: | |          : |     |  :     |:  | :| :|
Db         91 IACDLIKETKLSTEPSPGFSNYSEIAKFEKSVPDHCELVDDSSPESEPVDLFSDDSIPEV 150

Qy         90 PR------------------------------GPLPAAPPVAPERQPSW---------- 108
              |:                                |     |     ||:
Db        151 PQTQEEAVMLMKESLTEVSETVTQHKHKERLSASPQEVGKPYLESFQPNLHITKDAASNE 210

Qy        109 ----------------------------------------DPSPVSSTVPAPSPLSAA 126
                                                       | ||:      |: :||
Db        211 IPTLTKKETISLQMEEFNTAIYSNDDLLSSKEDKMKESETFSDSSPIEIIDEFPTFVSAK 270

Qy        127 AVSP------------------------------------------------------- 130
                ||
Db        271 DDSPKEYTDLEVSNKSEIANVQSGANSLPCSELPCDLSFKNTYPKDEAHVSDEFSKSRSS 330

Qy        131 ---------------------------------SKLPEDDEPPARPPPPPPASVSPQAE 156
                                                |||  | |    |
Db        331 VFKVPLLLLPNVSALESQIEMGNIVKPKVLTKEAEEKLPSDTEKEDR-------------- 376
```

```
Qy        157 PVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALAL 216
                :  |   |    |||||||||||||||||||||||||||||||||||||||||||||
Db        377 ---SLTAVLSAELNKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALAL 433

Qy        217 LSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELR 276
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        434 LSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELR 493

Qy        277 RLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG 336
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        494 RLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLG 553

Qy        337 LANKNVKDAMAKIQAKIPGLKRKAE 361
              ||||:|||||||||||||||||||
Db        554 LANKSVKDAMAKIQAKIPGLKRKAE 578
```

```
RESULT 11
Q7YRW9
ID   Q7YRW9      PRELIMINARY;      PRT;   184 AA.
AC   Q7YRW9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4w (Fragment).
GN   RTN4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22715887; PubMed=12832288;
RA   Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
RT   "A reticular rhapsody: phylogenic evolution and nomenclature of the
RT   RTN/Nogo gene family.";
RL   FASEB J. 17:1238-1247(2003).
DR   EMBL; AY164744; AAP47319.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   184 AA;  20671 MW;  DE990E03BBAF84A1 CRC64;

  Query Match             48.7%;  Score 901;  DB 6;  Length 184;
  Best Local Similarity   98.9%;  Pred. No. 3.9e-30;
  Matches  182;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        178 LYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSD 237
              |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LYWRDIKKTGVVFGXXLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSD 60

Qy        238 EGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVF 297
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 EGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVF 120

Qy        298 TYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLK 357
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 TYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLK 180

Qy        358 RKAE 361
              ||||
Db        181 RKAE 184


RESULT 12
Q7T224
ID   Q7T224      PRELIMINARY;      PRT;   199 AA.
AC   Q7T224;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   RTN4-C.
GN   RTN4.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22715887; PubMed=12832288;
RA   Oertle T., Klinger M., Stuermer C.A., Schwab M.E.;
RT   "A reticular rhapsody: phylogenic evolution and nomenclature of the
RT   RTN/Nogo gene family.";
RL   FASEB J. 17:1238-1247(2003).
DR   EMBL; AY164737; AAP47312.1; -.
SQ   SEQUENCE   199 AA;  22293 MW;  07CF4E4EF2723251 CRC64;

  Query Match             47.7%;  Score 883.2;  DB 13;  Length 199;
  Best Local Similarity   79.2%;  Pred. No. 2.7e-29;
  Matches  179;  Conservative    8;  Mismatches   11;  Indels   28;  Gaps    2;

Qy        136 DDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTSVVDLLYWRDIKKTGVVFGASLF 195
              | :|                    |              ||||||||||||||||||||
Db          2 DSQPSG----------------W-----------KDKVVDLLYWRDIKKTGVVFGASLF 33

Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEE 255
              ||||||||||||||||||||||||||||||||||||||||||||||||||:||:||:
Db         34 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESDVAVSED 93

Qy        256 LVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALI 315
              |:||||:  |||:| |:||||||||||||||||||||||||||||||||||||||||||
Db         94 LIQKYSSVVLGHINGTVKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALI 153

Qy        316 SLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||| |||||||||||||||||||| |
Db        154 SLFSVPVIYERHQAQIDHYLGLVNKNVKDAMAKIQAKIPGLKRKTE 199


RESULT 13
Q9GM33
ID   Q9GM33      PRELIMINARY;      PRT;   179 AA.
AC   Q9GM33;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
```

```
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA   Suzuki Y., Sugano S., Hashimoto K.;
RT   "Isolation of full-length cDNA clones from macaque brain cDNA
RT   libraries.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB049853; BAB16739.1; -.
DR   GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   179 AA;  19949 MW;  5F8CD4383FEE9E02 CRC64;

  Query Match           46.9%;  Score 867;  DB 6;  Length 179;
  Best Local Similarity  98.3%;  Pred. No. 9.6e-29;
  Matches  176;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        183 IKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPF 242
              :|||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db          1 MKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYRGVIQAIQKSDEGHPF 60

Qy        243 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGA 302
              |||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db         61 RAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLRFAVLMWVFTYVGA 120

Qy        303 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 179


RESULT 14
Q8K0T0
ID   Q8K0T0      PRELIMINARY;      PRT;   780 AA.
AC   Q8K0T0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to reticulon 1 (Hypothetical protein).
GN   RTN1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
```

RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC030455; AAH30455.1; -.
DR EMBL; BC053926; AAH53926.1; -.
DR MGD; MGI:1933947; Rtn1.
DR GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR InterPro; IPR001951; Histone_H4.
DR InterPro; IPR003388; Reticulon.
DR Pfam; PF02453; Reticulon; 1.
DR PROSITE; PS00047; HISTONE_H4; 1.
DR PROSITE; PS50845; RETICULON; 1.
KW Hypothetical protein.
SQ SEQUENCE 780 AA; 83572 MW; 29B47A58FC2F2027 CRC64;

Query Match 44.9%; Score 829.8; DB 11; Length 780;
Best Local Similarity 28.2%; Pred. No. 1.5e-25;
Matches 198; Conservative 60; Mismatches 91; Indels 352; Gaps 19;

```
Qy          4 LDQSPL----------------------------------------VSSSDSPPRPQ---- 20
              || ||                                          :::|:||  |:
Db         89 LDHSPSSTLKDGEGACYTSLISDVCYPPREDSAYFTGILQKENGHITTSESPEEPETPGP 148

Qy         21 -------------------------------------------PAFKYQFVREPEDEEEE 37
                                                         |:||  :  |:: : :
Db        149 SLPEVPGMEPQGLLSSDSGIEMTPAESTEVNKILADPLDQMKAEAYKYIDITRPQEAKGQ 208

Qy         38 EEEEEEDEDEDLE----ELEVLERKPAAGLSAAPVPTAPAAG-------------APLMD 80
              ||:    ||:||:    : ||  :  | |:  ||   ||  |             || :|
```

```
Db        209 EEQHPGLEDKDLDFKDKDTEVSTK--AEGV-RAPNQPAPVEGKLIKDHLFEESTFAPYID 265

Qy         81 FGND-----------------FVPPAPRGPLPAAPPVAPERQ-----PS---------- 107
              :|                    ||     : |     ||:|     ||
Db        266 ELSDEQHRVSLVTAPVKITLTEIEPPL----MTATQETIPEKQDLCLKPSPDTVPTVTVS 321

Qy        108 --WDPSPVSSTVPA---------------------------------------------- 119
                | || | | |:
Db        322 EPEDDSPGSVTPPSSGTEPSAAESQGKGSVSEDELIAAIKEAKGLSYETTESPRPVGQVA 381

Qy        120 --------------PSPL---------------------------------------- 123
                           ||||
Db        382 DKPKTKTRSGLPTIPSPLDQEASSAESGDSEIELVSEDPMASEDALPSGYVSFGHVSGPP 441

Qy        124 --------------------------------SAAAVSPSKLPEDDEPPARP------- 143
                                              ||:  || :  | | || :|
Db        442 PSPASPSIQYSILREEREAELDSELIIESCDASSASEESPKR--EQDSPPMKPGALDAIR 499

Qy        144 --------------------PPP-----PPASVSPQAEP-------VWTPP--------- 162
                                 | |     | |::  : ||       |  ||
Db        500 EETGSRATEERAPSHQGPVEPDPMLSFAPAAALQSRPEPSSGDGASVPEPPRSQQQKPEE 559

Qy        163 -------------APAPAA----PP-----STSVVDLLYWRDIKKTGVVFGASLFLLLSL 200
                          | |      ||         :||||||||:||:|||: | || ||
Db        560 EAVSSSQSPTATEIPGPLGSGLMPPLPFFNKQKAIDLLYWRDIKQTGIVFGSFLLLLFSL 619

Qy        201 TVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKY 260
              | ||:||| ||:||| || |||||||| |:||:||:||||||:|||| |: :|:| :|||
Db        620 TQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKY 679

Qy        261 SNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSV 320
              ::    :|| |:||||||||| ||||||||||||: ||||||||||||::|::|:|::
Db        680 TDCLQLYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMFTL 739

Qy        321 PVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              ||:| :||||:| ||||   ::   :|||||||| || ||
Db        740 PVVYVKHQAQVDQYLGLVRTHINTVVAKIQAKIPGAKRHAE 780
```

RESULT 15
Q8K4S4
ID   Q8K4S4   PRELIMINARY;   PRT;   780 AA.
AC   Q8K4S4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Reticulon 1A.
GN   RTN1 OR RTN-1A.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ICR; TISSUE=Brain;
RA   Hirata T., Nomura T., Takagi Y., Sato Y., Tomioka N., Fujisawa H.,
RA   Osumi N.;

```
RT   "Mosaic development of the olfactory cortex with Pax6-dependent and -
RT   independent components.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB074899; BAB96551.1; -.
DR   MGD; MGI:1933947; Rtn1.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   InterPro; IPR001951; Histone_H4.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS00047; HISTONE_H4; 1.
DR   PROSITE; PS50845; RETICULON; 1.
SQ   SEQUENCE   780 AA;  83504 MW;  545F5638C576A069 CRC64;

  Query Match             44.6%;  Score 825.8;  DB 11;  Length 780;
  Best Local Similarity   28.2%;  Pred. No. 2.3e-25;
  Matches  198;  Conservative   59;  Mismatches   92;  Indels  352;  Gaps   19;

Qy          4 LDQSPL----------------------------------------VSSSDSPPRPQ---- 20
              || ||                                        :::|:||  |:
Db         89 LDHSPSSTLKDGEGACYTSLISDVCYPPREDSAYFTGILQKENGHITTSESPEEPETPGP 148

Qy         21 ------------------------------------------PAFKYQFVREPEDEEEE 37
                                                        |:||  :  |:: : :
Db        149 SLPEVPGMEPQGLLSSDSGIEMTPAESTEVNKILADPLDQMKAEAYKYIDITRPQEAKGQ 208

Qy         38 EEEEEEDEDEDLE----ELEVLERKPAAGLSAAPVPTAPAAG-------------APLMD 80
              ||:    ||:||:      ||  :  | |:  ||   ||  |             || :|
Db        209 EEQHPGLEDKDLDFKDKGTEVSTK--AEGV-RAPNQPAPVEGKLIKDHLFEESTFAPYID 265

Qy         81 FGND-----------------FVPPAPRGPLPAAPPVAPERQ-----PS---------- 107
              :|                    ||     : |     ||:|     ||
Db        266 ELSDEQHRVSLVTAPVKITLTEIEPPL----MTATQETIPEKQDLCLKPSPDTVPTVTVS 321

Qy        108 --WDPSPVSSTVPA------------------------------------------------ 119
                | || | | |:
Db        322 EPEDDSPGSVTPPSSGTEPSAAESQGKGSVSEDELIAAIKEAKGLSYETTESPRPVGQVA 381

Qy        120 --------------PSPL---------------------------------------- 123
                           ||||
Db        382 DKPKTKTRSGLPTIPSPLDQEASSAESGDSEIELVSEDPMASEDALPSGYVSFGHVSGPP 441

Qy        124 ---------------------------------SAAAVSPSKLPEDDEPPARP------- 143
                                              ||:  || :  | | || :|
Db        442 PSPASPSIQYSILREEREAELDSELIIESCDASSASEESPKR--EQDSPPMKPGALDAIR 499

Qy        144 --------------------PPP-----PPASVSPQAEP-------VWTPP--------- 162
                               | |     | |::  : ||       |  ||
Db        500 EETGSRATEERAPSHQGPVEPDPMLSFAPAAALQSRPEPSSGDGASVPEPPRSQQQKPEE 559

Qy        163 -------------APAPAA----PP-----STSVVDLLYWRDIKKTGVVFGASLFLLLSL 200
                          | |      ||         :|||||||||:||:|||: | || ||
Db        560 EAVSSSQSPTATEIPGPLGSGLMPPLPFFNKQKAIDLLYWRDIKQTGIVFGSFLLLLFSL 619

Qy        201 TVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKY 260
              | ||:||| ||:||| || |||||||| |:||:||:|||||:|||| |: :|:| :|||
Db        620 TQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDEGHPFKAYLELEITLSQEQIQKY 679
```

```
Qy        261 SNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSV 320
              ::    :|| |:||||||||| |||||||||||||: ||||||||||||||::|::|:|::
Db        680 TDCLQLYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLTYVGALFNGLTLLLMAVVSMFTL 739

Qy        321 PVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              ||:| :||||:| ||||   ::   :|||||||| || ||
Db        740 PVVYVKHQAQVDQYLGLVRTHINTVVAKIQAKIPGAKRHAE 780
```

Search completed: September 29, 2004, 18:54:43
Job time : 80.6061 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 29, 2004, 18:21:11 ; Search time 15.0208 Seconds
(without alignments)
1251.419 Million cell updates/sec

Title: SEQ29_1-172_990-1178
Perfect score: 1850
Sequence: 1 MEDLDQSPLVSSSDSPPRPQ.....VKDAMAKIQAKIPGLKRKAE 361

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.1

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1756.9	95.0	1192	1	RTN4_HUMAN	Q9nqc3 homo sapien
2	1478.8	79.9	1163	1	RTN4_RAT	Q9jk11 rattus norv
3	911	49.2	199	1	RTN4_MOUSE	Q99p72 mus musculu
4	827.7	44.7	776	1	RTN1_HUMAN	Q16799 homo sapien
5	818.1	44.2	777	1	RTN1_RAT	Q64548 rattus norv
6	651.1	35.2	237	1	RTN3_MOUSE	Q9es97 mus musculu
7	644.2	34.8	236	1	RTN3_HUMAN	O95197 homo sapien
8	592.9	32.0	545	1	RTN2_HUMAN	O75298 homo sapien
9	571.9	30.9	471	1	RTN2_MOUSE	O70622 mus musculu
10	267.4	14.5	555	1	GP1_CHLRE	Q9fpq6 chlamydomon
11	266.4	14.4	1059	1	CAPU_DROME	Q24120 drosophila
12	266	14.4	1206	1	FM14_MOUSE	Q05859 mus musculu
13	266	14.4	1468	1	FMN1_MOUSE	Q05860 mus musculu
14	259.8	14.0	2517	1	NCR2_HUMAN	Q9y618 h nuclear r
15	259.7	14.0	1248	1	DIA1_HUMAN	O60610 homo sapien
16	257.1	13.9	3511	1	MY15_MOUSE	Q9qzz4 mus musculu
17	256.5	13.9	1255	1	DIA1_MOUSE	O08808 mus musculu

18	255.8	13.8	1157	1	BBC1_YEAST	P47068 saccharomyc
19	255.5	13.8	3530	1	MY15_HUMAN	Q9ukn7 homo sapien
20	253.8	13.7	1132	1	BAT3_HUMAN	P46379 homo sapien
21	252.9	13.7	2715	1	MLL4_HUMAN	Q9umn6 homo sapien
22	248.2	13.4	501	1	IRX3_HUMAN	P78415 homo sapien
23	247.3	13.4	980	1	FTKL_DEIRA	Q9rxb5 deinococcus
24	247	13.4	2090	1	N214_HUMAN	P35658 homo sapien
25	245.1	13.2	1300	1	SAL3_HUMAN	Q9bxa9 homo sapien
26	244.8	13.2	1375	1	BNR1_YEAST	P40450 saccharomyc
27	244.5	13.2	933	1	PRGR_HUMAN	P06401 homo sapien
28	244	13.2	1790	1	SEPA_EMENI	P78621 emericella
29	243.6	13.2	1719	1	PRD2_HUMAN	Q13029 homo sapien
30	243.3	13.2	786	1	PRGR_CHICK	P07812 gallus gall
31	243.2	13.1	507	1	IRX3_MOUSE	P81067 mus musculu
32	239.2	12.9	449	1	APG_BRANA	P40603 brassica na
33	239.1	12.9	930	1	PRGR_RABIT	P06186 oryctolagus
34	238.3	12.9	3149	1	TEGU_EBV	P03186 epstein-bar
35	236.9	12.8	980	1	RIN3_MOUSE	P59729 mus musculu
36	235.2	12.7	1953	1	BNI1_YEAST	P41832 saccharomyc
37	233.6	12.6	802	1	ENAH_MOUSE	Q03173 mus musculu
38	233.5	12.6	909	1	CNG4_HUMAN	Q14028 homo sapien
39	232.9	12.6	534	1	APG_ARATH	P40602 arabidopsis
40	232.5	12.6	1332	1	M4K6_HUMAN	Q8n4c8 homo sapien
41	231.2	12.5	1567	1	FMN2_MOUSE	Q9jl04 mus musculu
42	231.1	12.5	465	1	FXD1_HUMAN	Q16676 homo sapien
43	231.1	12.5	1362	1	BRD4_HUMAN	O60885 homo sapien
44	231	12.5	1164	1	FHOS_HUMAN	Q9y613 homo sapien
45	231	12.5	2774	1	MAPA_RAT	P34926 rattus norv

ALIGNMENTS

```
RESULT 1
RTN4_HUMAN
ID   RTN4_HUMAN     STANDARD;      PRT;  1192 AA.
AC   Q9NQC3; O94962; Q9BXG5; Q9H212; Q9H3I3; Q9UQ42; Q9Y293; Q9Y2Y7;
AC   Q9Y5U6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DE   (Neuroendocrine-specific protein) (NSP) (Neuroendocrine specific
DE   protein C homolog) (RTN-x) (Reticulon 5) (My043 protein).
GN   RTN4 OR NOGO OR ASY OR KIAA0886.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX   MEDLINE=20129242; PubMed=10667780;
RA   Prinjha R., Moore S.E., Vinson M., Blake S., Morrow R., Christie G.,
RA   Michalovich D., Simmons D.L., Walsh F.S.;
RT   "Inhibitor of neurite outgrowth in humans.";
RL   Nature 403:383-384(2000).
RN   [2]
```

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=21010696; PubMed=11126360;
RA Tagami S., Eguchi Y., Kinoshita M., Takeda M., Tsujimoto Y.;
RT "A novel protein, RTN-XS, interacts with both Bcl-XL and Bcl-2 on
RT endoplasmic reticulum and reduces their anti-apoptotic activity.";
RL Oncogene 19:5736-5746(2000).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX MEDLINE=20237542; PubMed=10773680;
RA Yang J., Yu L., Bi A.D., Zhao S.-Y.;
RT "Assignment of the human reticulon 4 gene (RTN4) to chromosome
RT 2p14-->2p13 by radiation hybrid mapping.";
RL Cytogenet. Cell Genet. 88:101-102(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RA Jin W.-L., Ju G.;
RT "Developmentally-regulated alternative splicing in a novel Nogo-A.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Placenta, and Skeletal muscle;
RA Ito T., Schwartz S.M.;
RT "Cloning of a member of the reticulon gene family in human.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fibroblast;
RA Yutsudo M.;
RT "Isolation of a cell death-inducing gene.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Pituitary;
RA Song H., Peng Y., Zhou J., Huang Q., Dai M., Mao Y.M., Yu Y., Xu X.,
RA Luo B., Hu R., Chen J.;
RT "Human neuroendocrine-specific protein C (NSP) homolog gene.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel human cDNA clone with function of inhibiting cancer cell
RT growth.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN [9]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosawa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [10]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Brain, Ovary, Pancreas, Placenta, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [11]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.-H., Ye M., Wu X.-Y., Ren S.-X., Zhao M., Zhao C.-J., Fu G.,
RA Shen Y., Fan H.-Y., Lu G., Zhong M., Xu X.-R., Han Z.-G., Zhang J.-W.,
RA Tao J., Huang Q.-H., Zhou J., Hu G.-X., Gu J., Chen S.-J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells.";
RL Genome Res. 10:1546-1560(2000).
RN [12]
RP SEQUENCE OF 482-1192 FROM N.A. (ISOFORM 1/4).
RC TISSUE=Brain;
RA Mao Y.M., Xie Y., Zheng Z.H.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN [13]
RP SEQUENCE OF 186-1192 FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN [14]
RP TOPOLOGY.
RC TISSUE=Brain;
RX MEDLINE=20129259; PubMed=10667797;
RA GrandPre T., Nakamura F., Vartanian T., Strittmatter S.M.;
RT "Identification of the Nogo inhibitor of axon regeneration as a
RT Reticulon protein.";
RL Nature 403:439-444(2000).
RN [15]
RP FUNCTION.
RC TISSUE=Brain;
RX MEDLINE=21069055; PubMed=11201742;
RA Fournier A.E., Grandpre T., Strittmatter S.M.;
RT "Identification of a receptor mediating Nogo-66 inhibition of axonal

```
RT   regeneration.";
RL   Nature 409:341-346(2001).
RN   [16]
RP   REVIEW.
RX   MEDLINE=21888956; PubMed=11891768;
RA   Ng C.E.L., Tang B.L.;
RT   "Nogos and the Nogo-66 receptor: factors inhibiting CNS neuron
RT   regeneration.";
RL   J. Neurosci. Res. 67:559-565(2002).
CC   -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC       block the regeneration of the nervous central system in adults.
CC       Isoform 2 reduces the anti-apoptotic activity of Bcl-xl and Bcl-2.
CC       This is likely consecutive to their change in subcellular
CC       location, from the mitochondria to the endoplasmic reticulum,
CC       after binding and sequestration.
CC   -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum. Anchored to the membrane of the endoplasmic reticulum
CC       through 2 putative transmembrane domains.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1; Synonyms=RTN 4A, Nogo-A, RTN-xL;
CC         IsoId=Q9NQC3-1; Sequence=Displayed;
CC       Name=2; Synonyms=RTN 4B, Nogo-B, RTN-xS, Foocen-M;
CC         IsoId=Q9NQC3-2; Sequence=VSP_005655;
CC       Name=3; Synonyms=RTN 4C, Nogo-C, Foocen-S;
CC         IsoId=Q9NQC3-3; Sequence=VSP_005652, VSP_005653;
CC       Name=4;
CC         IsoId=Q9NQC3-4; Sequence=VSP_005654;
CC   -!- TISSUE SPECIFICITY: Isoform 1 is specifically expressed in brain
CC       and testis and weakly in heart and skeletal muscle. Isoform 2 is
CC       widely expressed excepted for the liver. Isoform 3 is expressed in
CC       brain, skeletal muscle and adipocytes. Isoform 4 is testis-
CC       specific.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   -!- CAUTION: Ref.11 sequence differs from that shown due to
CC       frameshifts in positions 1149 and 1156.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ251383; CAB99248.1; -.
DR   EMBL; AJ251384; CAB99249.1; -.
DR   EMBL; AJ251385; CAB99250.1; -.
DR   EMBL; AB040462; BAB18927.1; -.
DR   EMBL; AB040463; BAB18928.1; -.
DR   EMBL; AF148537; AAG12176.1; -.
DR   EMBL; AF148538; AAG12177.1; -.
DR   EMBL; AF087901; AAG12205.1; -.
DR   EMBL; AF320999; AAG40878.1; -.
DR   EMBL; AF132047; AAD31021.1; -.
DR   EMBL; AF132048; AAD31022.1; -.
```

DR EMBL; AB015639; BAA83712.1; -.
DR EMBL; AF077050; AAD27783.1; -.
DR EMBL; AF177332; AAG17976.1; -.
DR EMBL; AB020693; BAA74909.1; -.
DR EMBL; BC001035; AAH01035.1; -.
DR EMBL; BC007109; AAH07109.1; -.
DR EMBL; BC014366; AAH14366.1; -.

Query Match 95.0%; Score 1756.9; DB 1; Length 1192;
Best Local Similarity 30.3%; Pred. No. 2e-50;
Matches 361; Conservative 0; Mismatches 0; Indels 831; Gaps 1;

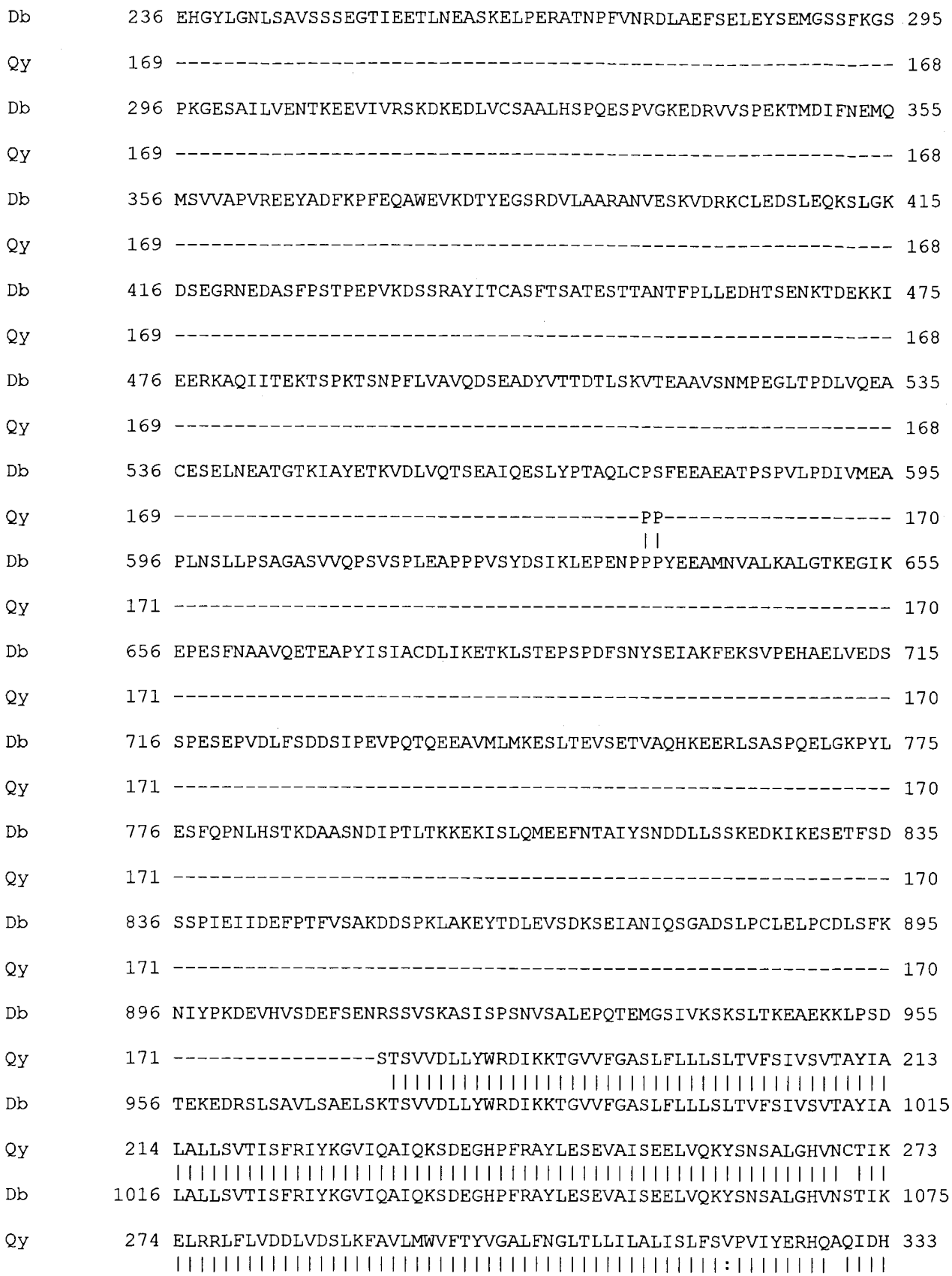
```
Qy          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEDLDQSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPA 60

Qy         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAP 120

Qy        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPS--------- 171
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTPAAPKRRG 180

Qy        172 ------------------------------------------------------------ 171

Db        181 SSGSVDETLFALPAASEPVIRSSAENMDLKEQPGNTISAGQEDFPSVLLETAASLPSLSP 240

Qy        172 ------------------------------------------------------------ 171

Db        241 LSAASFKEHEYLGNLSTVLPTEGTLQENVSEASKEVSEKAKTLLIDRDLTEFSELEYSEM 300

Qy        172 ------------------------------------------------------------ 171

Db        301 GSSFSVSPKAESAVIVANPREEIIVKNKDEEEKLVSNNILHNQQELPTALTKLVKEDEVV 360

Qy        172 ------------------------------------------------------------ 171

Db        361 SSEKAKDSFNEKRVAVEAPMREEYADFKPFERVWEVKDSKEDSDMLAAGGKIESNLESKV 420

Qy        172 ------------------------------------------------------------ 171

Db        421 DKKCFADSLEQTNHEKDSESSNDDTSFPSTPEGIKDRSGAYITCAPFNPAATESIATNIF 480

Qy        172 ------------------------------------------------------------ 171

Db        481 PLLGDPTSENKTDEKKIEEKKAQIVTEKNTSTKTSNPFLVAAQDSETDYVTTDNLTKVTE 540

Qy        172 ------------------------------------------------------------ 171

Db        541 EVVANMPEGLTPDLVQEACESELNEVTGTKIAYETKMDLVQTSEVMQESLYPAAQLCPSF 600

Qy        172 ------------------------------------------------------------ 171

Db        601 EESEATPSPVLPDIVMEAPLNSAVPSAGASVIQPSSSPLEASSVNYESIKHEPENPPPYE 660

Qy        172 ------------------------------------------------------------ 171
```

```
Db        661 EAMSVSLKKVSGIKEEIKEPENINAALQETEAPYISIACDLIKETKLSAEPAPDFSDYSE 720

Qy        172 ------------------------------------------------------------ 171

Db        721 MAKVEQPVPDHSELVEDSSPDSEPVDLFSDDSIPDVPQKQDETVMLVKESLTETSFESMI 780

Qy        172 ------------------------------------------------------------ 171

Db        781 EYENKEKLSALPPEGGKPYLESFKLSLDNTKDTLLPDEVSTLSKKEKIPLQMEELSTAVY 840

Qy        172 ------------------------------------------------------------ 171

Db        841 SNDDLFISKEAQIRETETFSDSSPIEIIDEFPTLISSKTDSFSKLAREYTDLEVSHKSEI 900

Qy        172 ------------------------------------------------------------ 171

Db        901 ANAPDGAGSLPCTELPHDLSLKNIQPKVEEKISFSDDFSKNGSATSKVLLLPPDVSALAT 960

Qy        172 ------------------------------------------TSVVDLLYWRDIKKTGVV 189
                                                       ||||||||||||||||||
Db        961 QAEIESIVKPKVLVKEAEKKLPSDTEKEDRSPSAIFSAELSKTSVVDLLYWRDIKKTGVV 1020

Qy        190 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 FGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESE 1080

Qy        250 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 VAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTL 1140

Qy        310 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 LILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKRKAE 1192
```

```
RESULT 2
RTN4_RAT
ID   RTN4_RAT       STANDARD;      PRT;  1163 AA.
AC   Q9JK11; Q9JK10; Q9R0D9; Q9WUE9; Q9WUF0;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein) (Foocen)
DE   (Glut4 vesicle 20 kDa protein).
GN   RTN4 OR NOGO.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 3), AND PARTIAL SEQUENCE.
RC   STRAIN=Sprague-Dawley; TISSUE=Adipocyte;
RX   MEDLINE=99249816; PubMed=10231557;
RA   Morris N.J., Ross S.A., Neveu J.M., Lane W.S., Lienhard G.E.;
RT   "Cloning and characterization of a 22 kDa protein from rat adipocytes:
RT   a new member of the reticulon family.";
```

```
RL   Biochim. Biophys. Acta 1450:68-76(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RX   MEDLINE=20129258; PubMed=10667796;
RA   Chen M.S., Huber A.B., Van der Haar M.E., Frank M., Schnell L.,
RA   Spillmann A.A., Christ F., Schwab M.E.;
RT   "Nogo-A is a myelin-associated neurite outgrowth inhibitor and an
RT   antigen for monoclonal antibody IN-1.";
RL   Nature 403:434-439(2000).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RC   STRAIN=Wistar Kyoto; TISSUE=Vascular smooth muscle;
RA   Ito T., Schwartz S.M.;
RT   "Cloning of a member of the reticulon gene family in rat: one of two
RT   minor splice variants.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION.
RX   MEDLINE=22033691; PubMed=12037567;
RA   GrandPre T., Li S., Strittmatter S.M.;
RT   "Nogo-66 receptor antagonist peptide promotes axonal regeneration.";
RL   Nature 417:547-551(2002).
CC   -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC       block the regeneration of the nervous central system in adults (By
CC       similarity).
CC   -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
CC       membrane of the endoplasmic reticulum through 2 putative
CC       transmembrane domains (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1; Synonyms=Nogo-A, NI-220-250;
CC         IsoId=Q9JK11-1; Sequence=Displayed;
CC       Name=2; Synonyms=Nogo-B, Foocen-M1;
CC         IsoId=Q9JK11-2; Sequence=VSP_005658;
CC       Name=3; Synonyms=Nogo-C, VP20;
CC         IsoId=Q9JK11-3; Sequence=VSP_005656, VSP_005657;
CC       Name=4; Synonyms=Foocen-M2;
CC         IsoId=Q9JK11-4; Sequence=VSP_005659;
CC   -!- TISSUE SPECIFICITY: Isoforms 1, 2 and 3 are present in optic
CC       nerve, spinal cord and cerebral cortex. Isoforms 1 and 2 are
CC       present in dorsal root ganglion, sciatic nerve and PC12 cells
CC       after longer exposure. Isoforms 2 and 3 are detected in kidney,
CC       cartilage, skin, lung and spleen. Isoform 3 is expressed at high
CC       level in skeletal muscle. In adult animals isoform 1 is expressed
CC       mainly in the nervous system.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; AF051335; AAF01564.1; -.
DR   EMBL; AJ242961; CAB71027.1; -.
DR   EMBL; AJ242962; CAB71028.1; -.
DR   EMBL; AJ242963; CAB71029.1; -.
DR   EMBL; AF132045; AAD31019.1; -.
DR   EMBL; AF132046; AAD31020.1; -.
DR   GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.
DR   GO; GO:0005635; C:nuclear membrane; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR   GO; GO:0030517; P:negative regulation of axon extension; ISS.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT   DOMAIN        1    989       CYTOPLASMIC (Potential).
FT   TRANSMEM    990   1010       POTENTIAL.
FT   DOMAIN     1011   1104       LUMENAL (Potential).
FT   TRANSMEM   1105   1125       POTENTIAL.
FT   DOMAIN     1126   1163       CYTOPLASMIC (Potential).
FT   DOMAIN      976   1163       RETICULON.
FT   DOMAIN       33     46       POLY-GLU.
FT   DOMAIN       73     76       POLY-ALA.
FT   DOMAIN      140    145       POLY-PRO.
FT   VARSPLIC      1    964       Missing (in isoform 3).
FT                               /FTId=VSP_005656.
FT   VARSPLIC    965    975       AVLSAELSKTS -> MDGQKKHWKDK (in isoform
FT                               3).
FT                               /FTId=VSP_005657.
FT   VARSPLIC    173    975       Missing (in isoform 2).
FT                               /FTId=VSP_005658.
FT   VARSPLIC    192    975       Missing (in isoform 4).
FT                               /FTId=VSP_005659.
FT   CONFLICT   1130   1131       MISSING (IN REF. 3; AAD31020).
SQ   SEQUENCE   1163 AA;  126386 MW;  8CB894B09E94F0B6 CRC64;

  Query Match             79.9%;  Score 1478.8;  DB 1;  Length 1163;
  Best Local Similarity   27.4%;  Pred. No. 3.5e-41;
  Matches  320;  Conservative   11;  Mismatches   25;  Indels  812;  Gaps    7;

Qy          1 MEDLDQSPLVSSS-DSPPRPQPAFKYQFVREPEDEE-EEEEEEEDEDEDLEELEVLERK 58
              |||:||| ||||| |||||| |||||||| |||||| |||||:||::|||||||||||||
Db          1 MEDIDQSSLVSSSTDSPPRPPPAFKYQFVTEPEDEEDEEEEEDEEEDDEDLEELEVLERK 60

Qy         59 PAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVP 118
              |||||||| ||  ||| |||:|| :| |||||||||||||| ||||||||: || :   |
Db         61 PAAGLSAAAVP--PAAAAPLLDFSSDSVPPAPRGPLPAAPPAAPERQPSWERSPAA---P 115

Qy        119 APSPLSAAAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPA------------- 165
              |||   |||| ||||||||||||||||||||  || |||   |  ||
Db        116 APSLPPAAAVLPSKLPEDDEPPARPPPPPPAGASPLAEPAAPPSTPAAPKRRGSGSVDET 175

Qy        166 ----PAA---------------------------------------------------- 168
                  |||
Db        176 LFALPAASEPVIPSSAEKIMDLMEQPGNTVSSGQEDFPSVLLETAASLPSLSPLSTVSFK 235

Qy        169 ------------------------------------------------------------ 168
```


Db 236 EHGYLGNLSAVSSSEGTIEETLNEASKELPERATNPFVNRDLAEFSELEYSEMGSSFKGS 295
Qy 169 ---------- 168
Db 296 PKGESAILVENTKEEVIVRSKDKEDLVCSAALHSPQESPVGKEDRVVSPEKTMDIFNEMQ 355
Qy 169 ---------- 168
Db 356 MSVVAPVREEYADFKPFEQAWEVKDTYEGSRDVLAARANVESKVDRKCLEDSLEQKSLGK 415
Qy 169 ---------- 168
Db 416 DSEGRNEDASFPSTPEPVKDSSRAYITCASFTSATESTTANTFPLLEDHTSENKTDEKKI 475
Qy 169 ---------- 168
Db 476 EERKAQIITEKTSPKTSNPFLVAVQDSEADYVTTDTLSKVTEAAVSNMPEGLTPDLVQEA 535
Qy 169 ---------- 168
Db 536 CESELNEATGTKIAYETKVDLVQTSEAIQESLYPTAQLCPSFEEAEATPSPVLPDIVMEA 595
Qy 169 ----------PP---------- 170
Db 596 PLNSLLPSAGASVVQPSVSPLEAPPPVSYDSIKLEPENPPPYEEAMNVALKALGTKEGIK 655
Qy 171 ---------- 170
Db 656 EPESFNAAVQETEAPYISIACDLIKETKLSTEPSPDFSNYSEIAKFEKSVPEHAELVEDS 715
Qy 171 ---------- 170
Db 716 SPESEPVDLFSDDSIPEVPQTQEEAVMLMKESLTEVSETVAQHKEERLSASPQELGKPYL 775
Qy 171 ---------- 170
Db 776 ESFQPNLHSTKDAASNDIPTLTKKEKISLQMEEFNTAIYSNDDLLSSKEDKIKESETFSD 835
Qy 171 ---------- 170
Db 836 SSPIEIIDEFPTFVSAKDDSPKLAKEYTDLEVSDKSEIANIQSGADSLPCLELPCDLSFK 895
Qy 171 ---------- 170
Db 896 NIYPKDEVHVSDEFSENRSSVSKASISPSNVSALEPQTEMGSIVKSKSLTKEAEKKLPSD 955
Qy 171 ----------STSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIA 213
Db 956 TEKEDRSLSAVLSAELSKTSVVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIA 1015
Qy 214 LALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIK 273
Db 1016 LALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIK 1075
Qy 274 ELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDH 333

```
Db       1076 ELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSIPVIYERHQVQIDH 1135

Qy        334 YLGLANKNVKDAMAKIQAKIPGLKRKAE 361
              |||||||:||||||||||||||||||:
Db       1136 YLGLANKSVKDAMAKIQAKIPGLKRKAD 1163


RESULT 3
RTN4_MOUSE
ID   RTN4_MOUSE     STANDARD;      PRT;   199 AA.
AC   Q99P72; Q9CTE3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Reticulon 4 (Neurite outgrowth inhibitor) (Nogo protein).
GN   RTN4 OR NOGO.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3T3-L1; TISSUE=Adipocyte;
RA   Coulson A.C., Craggs P.D., Morris N.J.;
RT   "Mouse vp20/RTN4C cDNA.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 170-199 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Potent neurite outgrowth inhibitor which may also help
CC       block the regeneration of the nervous central system in adults (By
CC       similarity).
CC   -!- SUBUNIT: Binds to RTN4R. Interacts with Bcl-xl and Bcl-2 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Anchored to the
CC       membrane of the endoplasmic reticulum through 2 putative
```

```
CC       transmembrane domains (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=1;
CC         Comment=A number of isoforms may be produced;
CC       Name=1;
CC         IsoId=Q99P72-1; Sequence=Displayed;
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF326337; AAK08076.1; -.
DR   EMBL; AK003859; -; NOT_ANNOTATED_CDS.
DR   MGD; MGI:1915835; Rtn4.
DR   GO; GO:0005783; C:endoplasmic reticulum; IDA.
DR   GO; GO:0030176; C:integral to endoplasmic reticulum membrane; ISS.
DR   GO; GO:0005635; C:nuclear membrane; ISS.
DR   GO; GO:0005515; F:protein binding; ISS.
DR   GO; GO:0019987; P:negative regulation of anti-apoptosis; ISS.
DR   GO; GO:0030517; P:negative regulation of axon extension; ISS.
DR   GO; GO:0007399; P:neurogenesis; IDA.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT   DOMAIN        1     25       CYTOPLASMIC (Potential).
FT   TRANSMEM     26     50       POTENTIAL.
FT   DOMAIN       51    137       LUMENAL (Potential).
FT   TRANSMEM    138    162       POTENTIAL.
FT   DOMAIN      163    199       CYTOPLASMIC (Potential).
FT   DOMAIN       12    199       RETICULON.
SQ   SEQUENCE   199 AA;  22466 MW;  07BE5D580059ED9C CRC64;

  Query Match             49.2%;  Score 911;  DB 1;  Length 199;
  Best Local Similarity   98.4%;  Pred. No. 4.3e-24;
  Matches  185;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        174 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 VVDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAI 71

Qy        234 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVL 293
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db         72 QKSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNSTIKELRRLFLVDDLVDSLKFAVL 131

Qy        294 MWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKI 353
              ||||||||||||||||||||||||:||||||||||||||||||:||||||||||||||
Db        132 MWVFTYVGALFNGLTLLILALISLFSIPVIYERHQAQIDHYLGLANKSVKDAMAKIQAKI 191

Qy        354 PGLKRKAE 361
              ||||||||
Db        192 PGLKRKAE 199
```

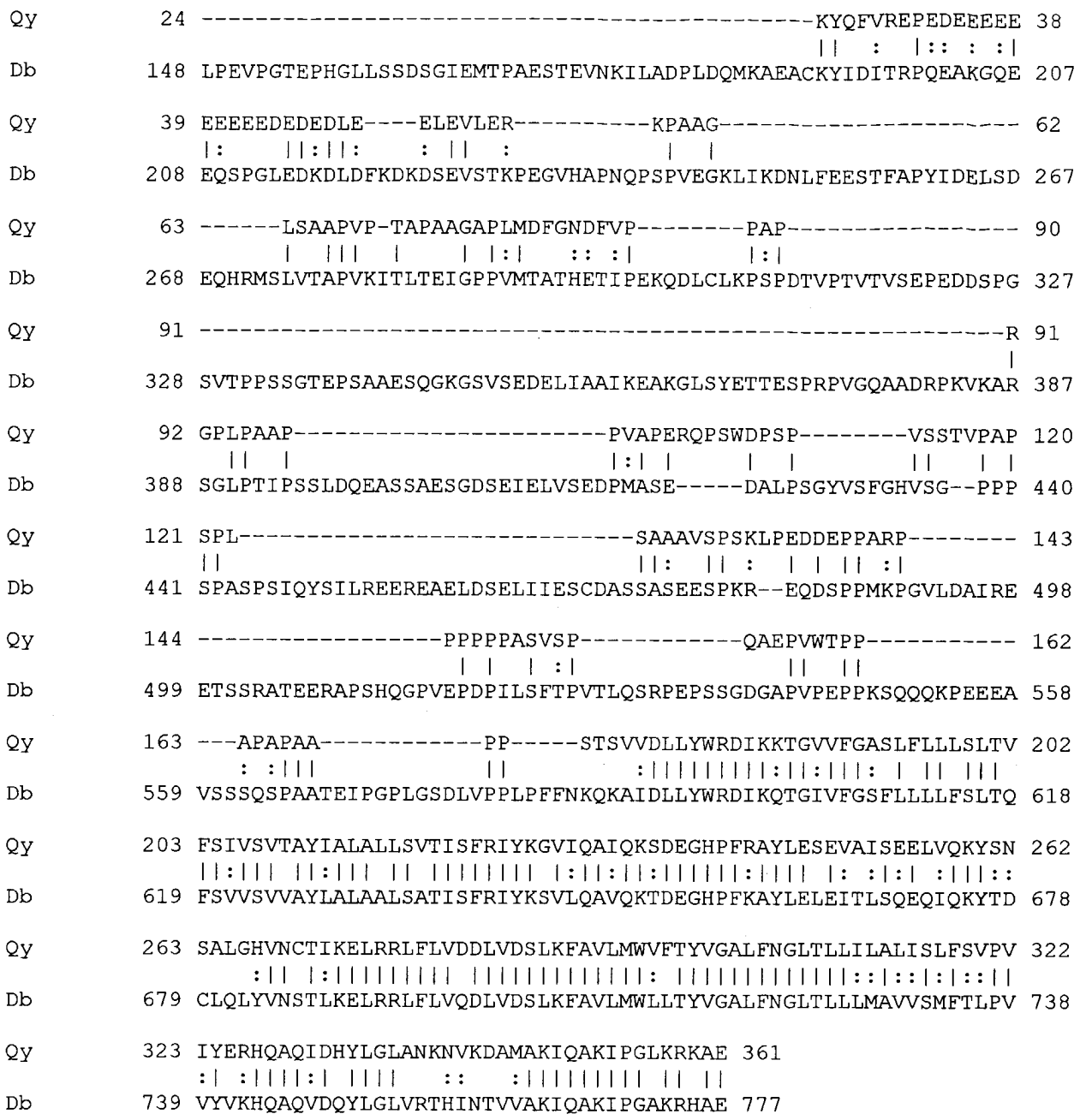
```
RESULT 4
RTN1_HUMAN
ID   RTN1_HUMAN     STANDARD;      PRT;   776 AA.
AC   Q16799; Q16800; Q16801;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon 1 (Neuroendocrine-specific protein).
GN   RTN1 OR NSP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS RTN1-A; RTN1-B AND RTN1-C).
RC   TISSUE=Lung carcinoma;
RX   MEDLINE=93293865; PubMed=7685762;
RA   Roebroek A.J.M., Van de Velde H.J.K., Van Bokhoven A., Broers J.L.V.,
RA   Ramaekers F.C.S., Van de Ven W.J.M.;
RT   "Cloning and expression of alternative transcripts of a novel
RT   neuroendocrine-specific gene and identification of its 135-kDa
RT   translational product.";
RL   J. Biol. Chem. 268:13439-13447(1993).
RN   [2]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=96429995; PubMed=8833145;
RA   Roebroek A.J.M., Ayoubi T.A.Y., Van de Velde H.J.K.,
RA   Schoenmakers E.F.P.M., Pauli I.G.L., Van de Ven W.J.M.;
RT   "Genomic organization of the human NSP gene, prototype of a novel gene
RT   family encoding reticulons.";
RL   Genomics 32:191-199(1996).
RN   [3]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=98228245; PubMed=9560466;
RA   Hens J., Nuydens R., Geerts H., Senden N.H., Van de Ven W.J.M.,
RA   Roebroek A.J., van de Velde H.J.K., Ramaekers F.C., Broers J.L.;
RT   "Neuronal differentiation is accompanied by NSP-C expression.";
RL   Cell Tissue Res. 292:229-237(1998).
CC   -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC       membrane trafficking in neuroendocrine cells.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=RTN1-A; Synonyms=NSP-A;
CC         IsoId=Q16799-1; Sequence=Displayed;
CC       Name=RTN1-B; Synonyms=NSP-B;
CC         IsoId=Q16799-2; Sequence=VSP_005644;
CC       Name=RTN1-C; Synonyms=NSP-C;
CC         IsoId=Q16799-3; Sequence=VSP_005645, VSP_005646;
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN NEURAL AND NEUROENDOCRINE TISSUES
CC       AND CELL CULTURES DERIVED THEREFROM. EXPRESSION OF ISOFORM RTN1-C
CC       IS STRONGLY CORRELATED WITH NEURONAL DIFFERENTIATION.
CC   -!- PTM: Isoforms RTN1-A and RTN1-B are phosphorylated.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10333; AAA59950.1; -.
DR   EMBL; L10334; AAA59951.1; -.
DR   EMBL; L10335; AAA59952.1; -.
DR   PIR; A46583; A46583.
DR   PIR; I60904; I60904.
DR   Genew; HGNC:10467; RTN1.
DR   MIM; 600865; -.
DR   GO; GO:0030176; C:integral to endoplasmic reticulum membrane; TAS.
DR   GO; GO:0004871; F:signal transducer activity; NAS.
DR   GO; GO:0030182; P:neuron differentiation; TAS.
DR   GO; GO:0007165; P:signal transduction; NAS.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW   Phosphorylation.
FT   TRANSMEM    603    623       POTENTIAL.
FT   TRANSMEM    726    746       POTENTIAL.
FT   DOMAIN      589    776       RETICULON.
FT   DOMAIN      609    612       POLY-LEU.
FT   VARSPLIC      1    420       Missing (in isoform RTN1-B).
FT                                /FTId=VSP_005644.
FT   VARSPLIC      1    568       Missing (in isoform RTN1-C).
FT                                /FTId=VSP_005645.
FT   VARSPLIC    569    588       GPGPLGPGAPPPLLFLNKQK -> MQATADSTKMDCVWSNW
FT                                KSQ (in isoform RTN1-C).
FT                                /FTId=VSP_005646.
SQ   SEQUENCE   776 AA;  83617 MW;  CA5B6232353096FE CRC64;

  Query Match             44.7%;  Score 827.7;  DB 1;  Length 776;
  Best Local Similarity   27.1%;  Pred. No. 6.8e-20;
  Matches  196;  Conservative   55;  Mismatches   89;  Indels  383;  Gaps   17;

Qy          6 QSPL---------------------------------------------------------- 9
              |||:
Db         70 QSPVAMETASTGVAGVSSAMDHTFSTTSKDGEGSCYTSLISDICYPPQEDSTYFTGILQK 129

Qy         10 ----VSSSDSPPR---PQP------------------------------------------- 21
                  |: |:||     | |
Db        130 ENGHVTISESPEELGTPGPSLPDVPGIESRGLFSSDSGIEMTPAESTEVNKILADPLDQM 189

Qy         22 ---AFKYQFVREPEDEEEEEEEEEDEDEDLE------------------------------ 50
                 |:||  :  ||: : :|:   | ||:||:
Db        190 KAEAYKYIDITRPEEVKHQEQHHPELEDKDLDFKNKDTDISIKPEGVREPDKPAPVEGKI 249

Qy         51 -----------------------------------ELE------------------- 53
                                                  |:|
Db        250 IKDHLLEESTFAPYIDDLSEEQRRAPQITTPVKITLTEIEPSVETTTQEKTPEKQDICLK 309
```

```
Qy         54 ----------------------------------------------VLERKPAAGLS-- 64
                                                           :   | | |||
Db        310 PSPDTVPTVTVSEPEDDSPGSITPPSSGTEPSAAESQGKGSISEDELITAIKEAKGLSYE 369

Qy         65 ----AAPV---------------PTAPAAGAPL------------------------- 78
                 ||               || |   :||
Db        370 TAENPRPVGQLADRPEVKARSGPPTIP---SPLDHEASSAESGDSEIELVSEDPMAAEDA 426

Qy         79 -----MDFGNDFVPPAPRGPLPAAPPV-----------------------------APE 103
                : ||:   ||    | ||:| :                                 :|:
Db        427 LPSGYVSFGHVGGPP----PSPASPSIQYSILREEREAELDSELIIESCDASSASEESPK 482

Qy        104 RQ---PSWDPSPVS------------------------------STVPAPSPLSAAAVSP 130
               |:   |   || :                            || | | |
Db        483 REQDSPPMKPSALDAIREETGVRAEERAPSRRGLAEPGSFLDYPSTEPQPGP-------- 534

Qy        131 SKLPEDD-----EPPARPPPPPPASVSPQAEPVWTPPAP-APAAPP------STSVVDLL 178
               :||  |     | |  |  |   | | |:      | |  | |||          :|||
Db        535 -ELPPGDGALEPETPMLPRKPEEDSSSNQSPAATKGPGPLGPGAPPPLLFLNKQKAIDLL 593

Qy        179 YWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDE 238
               ||||||:||:|||: | || ||| ||:||| ||:||| || |||||||| |:||:||:||
Db        594 YWRDIKQTGIVFGSFLLLLFSLTQFSVVSVVAYLALAALSATISFRIYKSVLQAVQKTDE 653

Qy        239 GHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFT 298
               ||||:|||| |: :|:| :|||::    :|| |:|||||||| ||||||||||||: |
Db        654 GHPFKAYLELEITLSQEQIQKYTDCLQFYVNSTLKELRRLFLVQDLVDSLKFAVLMWLLT 713

Qy        299 YVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGLKR 358
               |||||||||||::|::|:|::||:| :|||||| |||   ::   :|||||||| ||
Db        714 YVGALFNGLTLLLMAVVSMFTLPVVYVKHQAQIDQYLGLVRTHINAVVAKIQAKIPGAKR 773

Qy        359 KAE 361
               ||
Db        774 HAE 776
```

```
RESULT 5
RTN1_RAT
ID   RTN1_RAT       STANDARD;      PRT;   777 AA.
AC   Q64548; Q64547;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon 1 (Neuroendocrine-specific protein) (S-rex).
GN   RTN1 OR NSP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS RTN1-B AND RTN1-S).
RC   STRAIN=Wistar; TISSUE=Brain cortex;
RX   MEDLINE=96386034; PubMed=8793864;
RA   Baka I.D., Ninkina N.N., Pinon L.G.P., Adu J., Davies A.M.,
RA   Georgiev G.P., Buchman V.L.;
RT   "Intracellular compartmentalization of two differentially spliced s-
```

```
RT   rex/NSP mRNAs in neurons.";
RL   Mol. Cell. Neurosci. 7:289-303(1996).
CC   -!- FUNCTION: May be involved in neuroendocrine secretion or in
CC       membrane trafficking in neuroendocrine cells.
CC   -!- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane (By
CC       similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=RTN1-B; Synonyms=S-RexB;
CC         IsoId=Q64548-1; Sequence=Displayed;
CC       Name=RTN1-S; Synonyms=S-RexS;
CC         IsoId=Q64548-2; Sequence=VSP_005647, VSP_005648;
CC   -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CENTRAL AND
CC       PERIPHERAL NERVOUS SYSTEM OF NEWBORN AND ADULT RATS. LOW LEVELS
CC       HAVE BEEN ALSO DETECTED IN HEART, ADRENAL GLAND AND SPLEEN.
CC       EXPRESSION OF ISOFORM RTN1-B IS RESTRICTED TO PARTICULAR NEURONAL
CC       TYPES.
CC   -!- DEVELOPMENTAL STAGE: DETECTED ON EMBRYONIC DAY E10 IN THE
CC       HINDBRAIN AND IN E11 IN THE FOREBRAIN. DURING THE NEXT 3 EMBRYONIC
CC       DAYS THE LEVELS OF S-REXS INCREASES AND REMAINS STABLE AT E13 IN
CC       THE HINDBRAIN AND AT E14 IN THE FOREBRAIN. THE LEVELS OF S-REXB
CC       DOES NOT CHANGE AS SIGNIFICANTLY DURING DEVELOPMENT OF THE
CC       HINDBRAIN.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U17604; AAC53046.1; -.
DR   EMBL; U17603; AAC53045.1; -.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT   TRANSMEM    604    624       POTENTIAL.
FT   TRANSMEM    727    747       POTENTIAL.
FT   DOMAIN      590    777       RETICULON.
FT   DOMAIN      610    613       POLY-LEU.
FT   VARSPLIC      1    569       Missing (in isoform RTN1-S).
FT                               /FTId=VSP_005647.
FT   VARSPLIC    570    589       IPGPLGSDLVPPLPFFNKQK -> MQATADSTKMDCVWSNW
FT                               KSQ (in isoform RTN1-S).
FT                               /FTId=VSP_005648.
SQ   SEQUENCE   777 AA;  83001 MW;  AF7479C50F28D0AC CRC64;

  Query Match             44.2%;  Score 818.1;  DB 1;  Length 777;
  Best Local Similarity   28.2%;  Pred. No. 1.4e-19;
  Matches  197;  Conservative   56;  Mismatches   97;  Indels  349;  Gaps   19;

Qy          3 DLDQSP------------LVSSSDSPPRPQPAF-------------------------- 23
              |   ||            |:|    |||   |:
Db         88 DHSSSPTLKDGEGACYTSLISDICYPPREDSAYFTGILQKENGHITTSESPEELGTPGPS 147
```



RESULT 6
RTN3_MOUSE
ID   RTN3_MOUSE   STANDARD;   PRT;   237 AA.
AC   Q9ES97;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Reticulon protein 3.
GN   RTN3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Huang X., Zhou Y., Qiang H., Yuan J., Qiang B.;
RT   "Cloning and expression profile of a novel mouse cDNA encoding a human
RT   RTN3 homolog.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Eye;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF195940; AAG31360.1; -.
DR   EMBL; BC014697; AAH14697.1; -.
DR   MGD; MGI:1339970; Rtn3.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Transmembrane; Endoplasmic reticulum.
FT   TRANSMEM     69     89       POTENTIAL.
FT   TRANSMEM    167    187       POTENTIAL.
FT   DOMAIN       49    237       RETICULON.
SQ   SEQUENCE   237 AA;  25428 MW;  EB60A0A7AC45F0DE CRC64;

  Query Match             35.2%;  Score 651.1;  DB 1;  Length 237;
```

```
  Best Local Similarity   37.3%;  Pred. No. 3.1e-15;
  Matches  133;  Conservative   40;  Mismatches   55;  Indels  129;  Gaps    4;

Qy          6 QSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPAAGLSA 65
              ||| |||| |
Db          9 QSPSVSSSSS---------------------------------------------------- 18

Qy         66 APVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAPSPLSA 125
                     ||                                                 || |
Db         19 ---------GA-----------------------------------------EPSALGG 27

Qy        126 AAVSPSKLPEDDEPPARPPPPPPASVSPQAEPVWTPPAPAPAAPPSTSVVDLLYWRDIKK 185
                 ||                               |   | :   | :| ||::|||:||
Db         28 GGGSPGA--------------------------CPALGAKSCGSSCAVHDLIFWRDVKK 60

Qy        186 TGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRAY 245
              || ||| :| :|||| ||::|| :|: |||||||||||:|| ||||:|||:|||||:||
Db         61 TGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRVYKSVIQAVQKSEEGHPFKAY 120

Qy        246 LESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFN 305
              |: :: :| |    | |:|: |||  :| : |||||:|||||| || ||: |||||:||
Db        121 LDVDITLSSEAFHNYMNAAMVHVNKALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVFN 180

Qy        306 GLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-KRKAE 361
              |:|||||| : :||||::||::: |||||:|:|    |  : |||||:||: |:|||
Db        181 GITLLILAELLVFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKKAE 237


RESULT 7
RTN3_HUMAN
ID   RTN3_HUMAN     STANDARD;      PRT;   236 AA.
AC   O95197;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon protein 3 (Neuroendocrine-specific protein-like 2) (NSP-like
DE   protein II) (NSPLII).
GN   RTN3 OR NSPL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Retina;
RX   MEDLINE=99265974; PubMed=10331947;
RA   Moreira E.F., Jaworski C.J., Rodriguez I.R.;
RT   "Cloning of a novel member of the reticulon gene family (RTN3): gene
RT   structure and chromosomal localization to 11q13.";
RL   Genomics 58:73-81(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Huang X., Zhou Y., Du G., Yuan J., Qiang B.;
RT   "Cloning and expression analysis of a cDNA encoding a novel
RT   neuroendocrine-specific protein-like protein 1: NSPL1.";
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, Eye, and Lymph;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION IN
CC       BRAIN. THREE TIMES MORE ABUNDANT IN MACULA THAN IN PERIPHERAL
CC       RETINA.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF059524; AAC99319.1; -.
DR   EMBL; AF059529; AAD20951.1; -.
DR   EMBL; AF059525; AAD20951.1; JOINED.
DR   EMBL; AF059526; AAD20951.1; JOINED.
DR   EMBL; AF059527; AAD20951.1; JOINED.
DR   EMBL; AF059528; AAD20951.1; JOINED.
DR   EMBL; AF119297; AAD26810.1; -.
DR   EMBL; BC000634; AAH00634.1; -.
DR   EMBL; BC010556; AAH10556.1; -.
DR   EMBL; BC011394; AAH11394.1; -.
DR   EMBL; BC022993; AAH22993.1; -.
DR   Genew; HGNC:10469; RTN3.
DR   MIM; 604249; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Transmembrane; Endoplasmic reticulum.
FT   TRANSMEM     68     88       POTENTIAL.
```

```
FT   TRANSMEM    177    197       POTENTIAL.
FT   DOMAIN       48    236       RETICULON.
SQ   SEQUENCE   236 AA;  25609 MW;  DDC6A4544ABCDFB7 CRC64;

  Query Match             34.8%;  Score 644.2;  DB 1;  Length 236;
  Best Local Similarity   54.2%;  Pred. No. 5.1e-15;
  Matches  129;  Conservative   41;  Mismatches   50;  Indels   18;  Gaps    4;

Qy        138 EPPARPPPPPPASVSPQAEPVWTPPAP----APAAPP---------STSVVDLLYWRDIK 184
              || |       :| |  ||    | ||    :| | |         | :| ||::|||:|
Db          3 EPSAATQSHSISSSSFGAE----PSAPGGGGSPGACPALGTKSCSSSCAVHDLIFWRDVK 58

Qy        185 KTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHPFRA 244
              ||| ||| :| :|||| ||::|| :|: |||||||||||||| ||||:|||:|||||:|
Db         59 KTGFVFGTTLIMLLSLAAFSVISVVSYLILALLSVTISFRIYKSVIQAVQKSEEGHPFKA 118

Qy        245 YLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALF 304
              ||: :: :| |    | |:|: |:|  :| : |||||:|||||| || ||: |||||:|
Db        119 YLDVDITLSSEAFHNYMNAAMVHINRALKLIIRLFLVEDLVDSLKLAVFMWLMTYVGAVF 178

Qy        305 NGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPGL-KRKAE 361
              ||:|||||| : :||||::||::: |||||:|:|    |  : |||||:||: |:|||
Db        179 NGITLLILAELLIFSVPIVYEKYKTQIDHYVGIARDQTKSIVEKIQAKLPGIAKKKAE 236


RESULT 8
RTN2_HUMAN
ID   RTN2_HUMAN     STANDARD;      PRT;   545 AA.
AC   O75298; O60509;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE   protein 1) (NSPLI).
GN   RTN2 OR NSPL1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC   TISSUE=Lung carcinoma;
RX   MEDLINE=98360096; PubMed=9693037;
RA   Roebroek A.J.M., Contreras B., Pauli I.G.L., Van de Ven W.J.M.;
RT   "cDNA cloning, genomic organization, and expression of the human RTN2
RT   gene, a member of a gene family encoding reticulons.";
RL   Genomics 51:98-106(1998).
RN   [2]
RP   SEQUENCE OF 108-545 FROM N.A. (ISOFORM RTN2-B).
RC   TISSUE=Brain;
RX   MEDLINE=98191726; PubMed=9530622;
RA   Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT   "Molecular cloning of a novel mouse gene with predominant muscle and
RT   neural expression.";
RL   Mamm. Genome 9:274-282(1998).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
```

```
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=RTN2-A;
CC         IsoId=O75298-1; Sequence=Displayed;
CC         Note=Isoform RTN2-C is produced by alternative initiation at
CC         Met-341 of isoform RTN2-A;
CC       Name=RTN2-B;
CC         IsoId=O75298-2; Sequence=VSP_005649;
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, RTN2-A (shown here) and RTN2-C, are produced
CC         by alternative initiation at Met-1 and Met-341;
CC   -!- TISSUE SPECIFICITY: ISOFORM RTN2-C IS HIGHLY EXPRESSED IN SKELETAL
CC       MUSCLE.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF004222; AAC32542.1; -.
DR   EMBL; AF004223; AAC32543.1; -.
DR   EMBL; AF004224; AAC32544.1; -.
DR   EMBL; AF038540; AAC14910.1; -.
DR   Genew; HGNC:10468; RTN2.
DR   MIM; 603183; -.
DR   GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.
DR   GO; GO:0004871; F:signal transducer activity; NAS.
DR   GO; GO:0007165; P:signal transduction; NAS.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane;
KW   Alternative initiation.
FT   CHAIN         1    545       RETICULON PROTEIN 2, ISOFORM RTN2-A.
FT   CHAIN       341    545       RETICULON PROTEIN 2, ISOFORM RTN2-C.
FT   INIT_MET    341    341       FOR ISOFORM RTN2-C.
FT   TRANSMEM    368    388       POTENTIAL.
FT   TRANSMEM    463    483       POTENTIAL.
FT   DOMAIN      345    545       RETICULON.
FT   VARSPLIC    272    344       Missing (in isoform RTN2-B).
FT                                /FTId=VSP_005649.
SQ   SEQUENCE   545 AA;  59263 MW;  971FD2F909E1E9E6 CRC64;

  Query Match             32.0%;  Score 592.9;  DB 1;  Length 545;
  Best Local Similarity   27.3%;  Pred. No. 2e-12;
  Matches  153;  Conservative   56;  Mismatches  120;  Indels  231;  Gaps   16;

Qy          6 QSPLVSSSDSPPRPQPAFKYQFVREPEDEE-----------------EEEEEEEEDED- 46
              ::|  :||                 |:  |                   ||:|||   :|
Db         13 EAPSTASS---------------TPDSTEGGNDDSDFRELHTAREFSEEDEEETTSQDW 56

Qy         47 ------------------------------------------------EDLEELEVLE 56
                                                              : || :  |
```

```
Db         57 GTPRELTFSYIAFDGVVGSGGRRDSTARRPRPQGRSVSEPRDQHPQPSLGDSLESIPSLS 116

Qy         57 RKPAAGLSAAPVPTAPAAGAPLMD--FGNDFVPPAPRG-------------PLPAAPPVA 101
              : |  |    |  ||| :  || |     | :    ||             ||    |
Db        117 QSPEPGRRGDP-DTAPPSERPLEDLRLRLDHLGWVARGTGSGEDSSTSSSTPLEDEEPQE 175

Qy        102 PER-----------------QPSWDP-------SPVSSTVPAPSPLSAAAVSPSKLPEDD 137
              | |                 |||  |       || | |  | :|      |||:  : :
Db        176 PNRLETGEAGEELDLRLRLAQPS-SPEVLTPQLSPGSGTPQAGTP------SPSRSRDSN 228

Qy        138 EPPARP-------------------------------------------------------P 144
                | |
Db        229 SGPEEPLLEEEEKQWGPLEREPVRGQCLDSTDQLEFTVEPRLLGTAMEWLKTSLLLAVYK 288

Qy        145 PPPPASVSPQAEPVW----------TPP-----------APAPAAPPSTS--------V 174
                |   :||   |:|          |||           :|  :  || |        |
Db        289 TVPILELSP---PLWTAIGWVQRGPTPPTPVLRVLLKWAKSPRSSGVPSLSLGADMGSKV 345

Qy        175 VDLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQ 234
               |||||:| : :||||   :  || |  |||||| |::|| ||  ||| |:|: |:||:
Db        346 ADLLYWKDTRTSGVVFTGLMVSLLCLLHFSIVSVAAHLALLLLCGTISLRVYRKVLQAVH 405

Qy        235 KSDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLM 294
              : |  :||:|||: :: :: |  :: |:     |     :||  |||:|||||| |:|
Db        406 RGDGANPFQAYLDVDLTLTREQTERLSHQITSRVVSAATQLRHFFLVEDLVDSLKLALLF 465

Qy        295 WVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIP 354
              :: |:|||:|||||||| :| ||::|::| :|||||| |:||    :    |||:||||
Db        466 YILTFVGAIFNGLTLLILGVIGLFTIPLLYRQHQAQIDQYVGLVTNQLSHIKAKIRAKIP 525

Qy        355 GL-------------KRKAE 361
              |              | |||
Db        526 GTGALASAAAAVSGSKAKAE 545
```

```
RESULT 9
RTN2_MOUSE
ID   RTN2_MOUSE     STANDARD;      PRT;   471 AA.
AC   O70622; O70620;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Reticulon protein 2 (Neuroendocrine-specific protein-like 1) (NSP-like
DE   protein 1) (NSPLI).
GN   RTN2 OR NSPL1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND TISSUE SPECIFICITY.
RC   STRAIN=FVB/N, and 129/Sv; TISSUE=Cerebellum, and Skeletal muscle;
RX   MEDLINE=98191726; PubMed=9530622;
RA   Geisler J.G., Stubbs L.J., Wasserman W.W., Mucenski M.L.;
RT   "Molecular cloning of a novel mouse gene with predominant muscle and
RT   neural expression.";
RL   Mamm. Genome 9:274-282(1998).
```

```
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Retina;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
CC       (Potential).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=Brain;
CC         IsoId=O70622-1; Sequence=Displayed;
CC       Name=2; Synonyms=Muscle;
CC         IsoId=O70622-2; Sequence=VSP_005650, VSP_005651;
CC   -!- TISSUE SPECIFICITY: Expressed predominantly in neural and muscular
CC       tissues.
CC   -!- SIMILARITY: Contains 1 reticulon domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF038537; AAC14906.1; -.
DR   EMBL; AF038537; AAC14907.1; -.
DR   EMBL; AF038538; AAC14908.1; -.
DR   EMBL; AF038539; AAC14909.1; -.
DR   EMBL; AF093624; AAD13195.1; -.
DR   EMBL; BC031370; AAH31370.1; -.
DR   MGD; MGI:107612; Rtn2.
DR   InterPro; IPR003388; Reticulon.
DR   Pfam; PF02453; Reticulon; 1.
DR   PROSITE; PS50845; RETICULON; 1.
KW   Endoplasmic reticulum; Alternative splicing; Transmembrane.
FT   TRANSMEM    295    315       POTENTIAL.
FT   DOMAIN      272    471       RETICULON.
```

```
FT   VARSPLIC      1    267       Missing (in isoform 2).
FT                                /FTId=VSP_005650.
FT   VARSPLIC    268    271       PLLL -> MGSK (in isoform 2).
FT                                /FTId=VSP_005651.
SQ   SEQUENCE   471 AA;  51346 MW;  9BBD8F372CF63AD3 CRC64;

  Query Match             30.9%;  Score 571.9;  DB 1;  Length 471;
  Best Local Similarity   28.3%;  Pred. No. 7e-12;
  Matches  141;  Conservative   55;  Mismatches  121;  Indels  181;  Gaps   13;

Qy          6 QSPLVSSSDSPPRPQPAFKYQFVREPEDEE-----------------EEEEEEEEDED- 46
              ::|  :||                 |:  |                   |:||||   :|
Db         13 EAPSTASS----------------TPDSTEGGNDDSDFRELHTAREFSEDEEEETTSQDW 56

Qy         47 ------------------------------------------------EDLEELEVLE 56
                                                              : || :  |
Db         57 GTPRELTFSYIAFDGVVGSGGRRDSVVRRPRPQGRSVSEPRDPPQQSGLGDSLESIPSLS 116

Qy         57 RKPAAGLSAAPVPTAPAAGAPLMDFGNDFVPPAPRGPLPAAPPVAPER----QPSW---- 108
              : |  |    | |                |||| |       |:   |    |  |
Db        117 QSPEPGRRGDPDP----------------VPPAER-------PLEELRLRLDQLGWVVRS 153

Qy        109 ----DPSPVSSTVP----APSPLSAA--------------------AVSPSKLPEDDEP 139
                 : |  ||: |     |  | |:                    ::|   |    |
Db        154 AGSGEDSATSSSTPLENEEPDGLEASEAGEETNLELRLAQSLHLQLEVLTPQLSPSSGTP 213

Qy        140 PARPPPPPPASVS---PQAEPV-------WTPPAPAPAA--------------PPSTSVV 175
               |  | |  :  |   |  ||:       |      |                |   |
Db        214 QAHTPSPQRSQDSNSGPDDEPLLNVVEEHWRLLEQEPITAQCLDSTDQSEFMLEPLLLVA 273

Qy        176 DLLYWRDIKKTGVVFGASLFLLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQK 235
              |||||:| : :| ||   :  || |  |||||| |::||  |  ||| |:|: |:||: :
Db        274 DLLYWKDTRTSGAVFTGLMASLLCLLHFSIVSVAAHLALLGLCATISLRVYRKVLQAVHR 333

Qy        236 SDEGHPFRAYLESEVAISEELVQKYSNSALGHVNCTIKELRRLFLVDDLVDSLKFAVLMW 295
               |  :||:|||: :: :: |  :: |     ||  |  :||  |||:|||||| |:| :
Db        334 GDGTNPFQAYLDMDLTLTREQTERLSQQIASHVVSTATQLRHFFLVEDLVDSLKLALLFY 393

Qy        296 VFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAKIQAKIPG 355
              : |:|||:||||||:|| :::||:||::| :|||||| |:||    :    |||:|||||
Db        394 ILTFVGAIFNGLTLVILGVVALFTVPLLYRQHQAQIDQYVGLVTNQLSHIKAKIRAKIPG 453

Qy        356 L------------KRKAE 361
                          | |||
Db        454 TGTLAPTASVSGSKAKAE 471


RESULT 10
GP1_CHLRE
ID   GP1_CHLRE      STANDARD;      PRT;   555 AA.
AC   Q9FPQ6; Q03927;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Vegetative cell wall protein gp1 precursor (Hydroxyproline-rich
DE   glycoprotein 1).
GN   GP1.
```

```
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21159092; PubMed=11258910;
RA   Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
RA   Goodenough U.W.;
RT   "Glycosylated polyproline II rods-with-kinks as a structural motif in
RT   plant hydroxyproline-rich glycoproteins.";
RL   Biochemistry 40:2978-2987(2001).
RN   [2]
RP   PARTIAL PRELIMINARY SEQUENCE FROM N.A.
RX   MEDLINE=91017504; PubMed=1699225;
RA   Adair W.S., Apt K.E.;
RT   "Cell wall regeneration in Chlamydomonas: accumulation of mRNAs
RT   encoding cell wall hydroxyproline-rich glycoproteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
CC   -!- FUNCTION: Major component of the outer cell wall W6 (crystalline)
CC       layer.
CC   -!- SUBUNIT: Associates with GP2 and GP3.
CC   -!- PTM: N-glycosylated and O-glycosylated.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF309494; AAG45420.1; -.
DR   EMBL; M58496; AAA69706.1; ALT_SEQ.
DR   GlycoSuiteDB; Q9FPQ6; -.
DR   InterPro; IPR003882; Pistil_extensin.
DR   PRINTS; PR01218; PSTLEXTENSIN.
KW   Glycoprotein; Repeat; Signal.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30    555       VEGETATIVE CELL WALL PROTEIN GP1.
FT   DOMAIN       40    339       49 X 5 AA APPROXIMATE PPSPX REPEATS.
FT   DOMAIN      259    279       POLY-PRO.
FT   CARBOHYD    399    399       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    455    455       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    493    493       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   555 AA;  54219 MW;  6A584A90465502F5 CRC64;

  Query Match             14.5%;  Score 267.4;  DB 1;  Length 555;
  Best Local Similarity   16.5%;  Pred. No. 0.15;
  Matches   90;  Conservative   30;  Mismatches  100;  Indels  326;  Gaps   15;

Qy          7 SPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEEDEDEDLEELEVLERKPAAGLSAA 66
              :| :  | :|| | |                                          |     :
Db        143 APPLPPSPAPPSPSP---------------------------------------PVPPSPSP 165

Qy         67 PVPTAPAAGAPLMDFGNDFVPPAPRGPLPA------------APPVAPERQPSWDPSPVS 114
              ||| :||  :|     :  |||:|  | ||            |||| |   |   |||
```

```
Db        166 PVPPSPAPPSPTPPSPSPPVPPSPAPPSPAPPVPPSPAPPSPAPPVPPSPAPPSPPSPAP 225

Qy        115 STVPA-----------PSPLSAAAVSPSKLPEDDEPPARPPPP---------------- 146
               : |:           |||:  :   ||  |   :||| ||||
Db        226 PSPPSPAPPSPSPPAPPSPVPPSPAPPSPAPPSPKPPAPPPPPSPPPPPPPRPPFPANTP 285

Qy        147 ------------------------PPASVSPQAEPV---WTPPAPAPAAPPS-------- 171
                                      ||: | |   ||     ||:|||: |||
Db        286 MPPSPPSPPPSPAPPTPPTPPSPSPPSPVPPSPAPVPPSPAPPSPAPSPPPSPAPPTPSP 345

Qy        172 ----------------------------------TSVVDLLYWRD--------------- 182
                                               : |   | | |
Db        346 SPSPSPSPSPSPSPSPSPSPSPSPIPSPSPKPSPSPVAVKLVWADDAIAFDDLNGTSTRP 405

Qy        183 --------------IKKTG-----------------VVFGASLFLLLSLTVFSIVSVTAY 211
                           |  |                  ||          || |: :||
Db        406 GSASRMVGEPDIAGTKCKGNLKGWMPKPSRNPRWGQAVFSG------GRTVGSVANVTIR 459

Qy        212 IAL-----ALLSVTISFRIYKGVIQAIQKSDEGHPFRAYLESEVAISEELVQKYSNSALG 266
               :|      ||:  :|   :|
Db        460 VAFATEKPALIYSSIELVVYN------------------------------------- 480

Qy        267 HVNCTIKELRRLFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVI--Y 324
                                                 | ||:         ||:
Db        481 ---------------------------------------TGATLI--------RVPIAANV 494

Qy        325 ERHQAQIDHYL-------------------------------GLANKNVKDAMAKIQAKI 353
               | | :   :|                              |  ||  | ::  :
Db        495 TRSQIRCPGFLTYGTTPIAGYPTGIDATTWPNWKIAGVRINMGAGNKKPKTSIDAV---- 550

Qy        354 PGLKRK 359
               ||  |
Db        551 -GLNLK 555
```

```
RESULT 11
CAPU_DROME
ID   CAPU_DROME     STANDARD;      PRT;  1059 AA.
AC   Q24120; Q9VQV8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cappuccino protein.
GN   CAPU OR CG3399.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=96033799; PubMed=7590229;
RA   Emmons S., Phan H., Calley J., Chen W., James B., Manseau L.;
RT   "Cappuccino, a Drosophila maternal effect gene required for polarity
RT   of the egg and embryo, is related to the vertebrate limb deformity
RT   locus.";
```

RL Genes Dev. 9:2482-2494(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
CC subfamily.
CC ---

CC ---
DR EMBL; U34258; AAC46925.1; -.

```
DR   EMBL; AE003578; AAF51054.1; -.
DR   PIR; T13286; T13286.
DR   FlyBase; FBgn0000256; capu.
DR   GO; GO:0007304; P:eggshell formation (sensu Insecta); IMP.
DR   GO; GO:0007316; P:pole plasm RNA localization; IMP.
DR   InterPro; IPR000269; CuNH_oxidase.
DR   InterPro; IPR003104; FH2.
DR   InterPro; IPR001265; Formin.
DR   Pfam; PF02181; FH2; 1.
DR   PRINTS; PR00828; FORMIN.
DR   SMART; SM00498; FH2; 1.
KW   Developmental protein.
FT   DOMAIN      480    560       FH1 (PRO-RICH).
FT   DOMAIN      585   1021       FH2.
FT   CONFLICT    260    260       S -> C (IN REF. 1).
FT   CONFLICT    364    364       S -> T (IN REF. 1).
FT   CONFLICT    386    386       T -> S (IN REF. 1).
FT   CONFLICT    471    471       E -> K (IN REF. 1).
FT   CONFLICT    495    495       H -> P (IN REF. 1).
FT   CONFLICT    513    513       MISSING (IN REF. 1).
SQ   SEQUENCE   1059 AA;  113863 MW;  009B0E24F61B6EA5 CRC64;

  Query Match             14.4%;  Score 266.4;  DB 1;  Length 1059;
  Best Local Similarity   16.2%;  Pred. No. 0.75;
  Matches  119;  Conservative   46;  Mismatches  114;  Indels  456;  Gaps   27;

Qy          9 LVSSSDSPPR--------------PQPAFKYQF--------VREPEDEEEEEEEEE---- 42
              |:|| | |||                 |  |::        |:   | | : | |:
Db        398 LLSSVDEPPRRPSKRCVNLTELLNASEATVYEYNKTGAEGCVKSFTDAETQTESEDCEGT 457

Qy         43 -------------EDEDEDLEELEVLERKPAAGLSAAPVPTAP-----AAGAPLMDFGND 84
                           |   || |       || |    || |  |      |  ||
Db        458 CKCGQSSTKVSDNESAKEDGE-------KPHA---VAPPPPPPPPPLHAFVAP------- 500

Qy         85 FVPPAPRGPLPAAPPVAPERQPSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPAR-- 142
                || |  | |  ||:|    |   | |     | | | | :|  |        |||
Db        501 --PPPPPPPPPPPPLANYGAP---PPP-----PPPPPGSGSAPPP-------PPPAPIE 543

Qy        143 -----PPPPPPASVSPQAEPV-----------------------------------WT-- 160
                  |||||| | ||    :                                    ||
Db        544 GGGGIPPPPPPMSASPSKTTISPAPLPDPAEGNWFHRTNTMRKSAVNPPKPMRPLYWTRI 603

Qy        161 ----PPAP-----------------------------APAAPPSTSVVDLLYWRDIKKT 186
                  ||||                           | |||:|  :   | :|::|
Db        604 VTSAPPAPRPPSVANSTDSTENSGSSPDEPPAANGADAPPTAPPATKEI----WTEIEET 659

Qy        187 ----------------------------------------------GVVF---------- 190
                                                           |:::
Db        660 PLDNIDEFTELFSRQAIAPVSKPKELKVKRAKSIKVLDPERSRNVGIIWRSLHVPSSEIE 719

Qy        191 ----------------------------------GASLFL------LLSLTVFSI--- 205
                                               |  : |      || ::: |:
Db        720 HAIYHIDTSVVSLEALQHMSNIQATEDELQRIKEAAGGDIPLDHPEQFLLDISLISMASE 779

Qy        206 -VSVTAYIALALLSVTISFRIYKGVIQAIQ---------------------------- 234
               :|   : |    |||: ||  : | |  |
Db        780 RISCIVFQAEFEESVTLLFRKLETVSQLSQQLIESEDLKLVFSIILTLGNYMNGGNRQRG 839
```

```
Qy        235 ----------------KSDEGH---------------------------------------- 240
                              || | |
Db        840 QADGFNLDILGKLKDVKSKESHTTLLHFIVRTYIAQRRKEGVHPLEIRLPIPEPADVERA 899

Qy        241 ---------------------------------------PFRAYLESEVAISEELVQKYS 261
                                                    ||:: :|  |  ::: : |
Db        900 AQMDFEEVQQQIFDLNKKFLGCKRTTAKVLAASRPEIMEPFKSKMEEFVEGADKSMAKLH 959

Qy        262 NSALGHVNCTIKELRRLFL--------------------------------VDDLVDSLK 289
               |        : | | |||                                  :|  |  |
Db        960 QS--------LDECRDLFLETMRFYHFSPKACTLTLAQCTPDQFFEYWTNFTNDFKDIWK 1011

Qy        290 FAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVKDAMAK- 348
                     : :| | |               : :  ||||:     : :||   : |
Db       1012 --------KEITSLLNEL--------------MKKSKQAQIE-----SRRNVSTKVEKS 1043

Qy        349 --IQAKIPGLKRKAE 361
                 |  |   | |:::
Db       1044 GRISLKERMLMRRSK 1058
```

```
RESULT 12
FM14_MOUSE
ID   FM14_MOUSE     STANDARD;      PRT;  1206 AA.
AC   Q05859;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Formin 1 isoform IV (Limb deformity protein).
GN   FMN OR LD.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=92112033; PubMed=1339380;
RA   Grusby-Jackson L., Kuo A., Leder P.;
RT   "A variant limb deformity transcript expressed in the embryonic mouse
RT   limb defines a novel formin.";
RL   Genes Dev. 6:29-37(1992).
CC   -!- FUNCTION: Is important in the morphogenesis of limb and may have a
CC       function in differentiated cells or be involved in maintaining
CC       specific differentiated states.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC         Comment=Additional isoforms seem to exist;
CC       Name=IV;
CC         IsoId=Q05859-1; Sequence=Displayed;
CC       Name=IA;
CC         IsoId=Q05860-1; Sequence=External;
CC       Name=IB;
CC         IsoId=Q05860-2; Sequence=External;
CC       Name=II;
CC         IsoId=Q05860-3; Sequence=External;
```

```
CC       Name=III;
CC         IsoId=Q05860-4; Sequence=External;
CC   -!- TISSUE SPECIFICITY: It is found throughout the embryo but has a
CC       functional role only in the kidney and limb.
CC   -!- DEVELOPMENTAL STAGE: This is the isoform found in the apical
CC       ectodermal ridge and the mesenchymal compartment of the developing
CC       limb bud.
CC   -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC   -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC   -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
CC       subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X62379; CAA44244.1; -.
DR   PIR; S24407; S24407.
DR   MGD; MGI:101815; Fmn.
DR   InterPro; IPR003104; FH2.
DR   InterPro; IPR001265; Formin.
DR   Pfam; PF02181; FH2; 1.
DR   PRINTS; PR00828; FORMIN.
KW   Nuclear protein; Developmental protein; Alternative splicing;
KW   Phosphorylation; Coiled coil.
FT   DOMAIN      418    443       COILED COIL (POTENTIAL).
FT   DOMAIN      497    566       COILED COIL (POTENTIAL).
FT   DOMAIN      644    744       FH1 (PRO-RICH).
FT   DOMAIN      759   1164       FH2.
FT   DOMAIN     1043   1116       COILED COIL (POTENTIAL).
FT   DOMAIN      635    638       POLY-SER.
FT   DOMAIN      751    755       POLY-SER.
SQ   SEQUENCE   1206 AA;  133464 MW;  4DFB38CB52BD8EE7 CRC64;

  Query Match             14.4%;  Score 266;  DB 1;  Length 1206;
  Best Local Similarity   17.3%;  Pred. No. 1.1;
  Matches  102;  Conservative   40;  Mismatches   97;  Indels  350;  Gaps   24;

Qy          5 DQSPLVSSSDSPPR-----PQPAFKYQFVREPEDEEE------------------EEEEE 41
              ||||      :| |     | |  |        | ||                  : | :
Db        459 DQSPTEQDDRTPGRLQAVWPPPKTK--------DTEEKVGLKYTEAEYQAAILHLKREHK 510

Qy         42 EEDE-----------------------DEDLEEL-EVLERK------------------ 58
              || |                      :| :| | : ||::
Db        511 EEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKRREGCEEMRDVCISTDDDC 570

Qy         59 -PAA---------------------------------------------------------- 61
               | |
Db        571 SPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISLTQLSPSKDSKDIHAPFQT 630

Qy         62 --GLSA---------APVPTAPAAGAPLMDFGNDFVPPAP-----RGPLPAAPPVAPERQ 105
                | |:         || || |    ||       :|| |      |||| |||: |
```

```
Db        631 REGTSSSSQQKISPPAP-PTPPPLPPPL-------IPPPPPLPPGLGPLPPAPPIPP--- 679

Qy        106 PSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPA--------SVSPQAEP 157
                  |||   | | |       |: :|  | ||  ||||||         |  |   |
Db        680 ----VCPVSPPPPPPPP------PPTPVPPSDGPPPPPPPPPPLPNVLALPNSGGPPPPP 729

Qy        158 VWTPPAP--APAAPPSTSV--------------------VDLLYWRDIKKTGVVFGASLF 195
                 || |  ||  ||  |                     :  |||
Db        730 PPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKPLYWTR------------- 776

Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHP----------FRAY 245
                                            ||   || :  |           |
Db        777 ---------------------------------IQINDKSQDAAPTLWDSLEEPHIRDT 802

Qy        246 LESEVAISEELVQ-----------------------KYSNSA---LGHVNCTIKELRR 277
               | |   |::  |                       | | :    :  ::  :|::::
Db        803 SEFEYLFSKDTTQQKKKPLSEAYEKKNKVKKIIKLLDGKRSQTVGILISSLHLEMKDIQQ 862

Qy        278 -LFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG 336
              :| |||                            ::: | ::  :||
Db        863 AIFTVDD---------------------------SVVDLETLAALYE----------- 882

Qy        337 LANKNVKDAMAKIQ----------------------AKIPGLKRKAE 361
               |:  :| : ||:                      |:||    :|:
Db        883 --NRAQEDELTKIRKYYETSKEEDLKLLDKPEQFLHELAQIPNFAERAQ 929
```

```
RESULT 13
FMN1_MOUSE
ID   FMN1_MOUSE     STANDARD;      PRT;  1468 AA.
AC   Q05860;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Formin 1 isoforms I/II/III (Limb deformity protein).
GN   FMN OR LD.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Testis;
RX   MEDLINE=90363291; PubMed=2392150;
RA   Woychik R.P., Maas R.L., Zeller R., Vogt T.F., Leder P.;
RT   "'Formins': proteins deduced from the alternative transcripts of the
RT   limb deformity gene.";
RL   Nature 346:850-853(1990).
RN   [2]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=97224459; PubMed=9119367;
RA   Wang C.C., Chan D.C., Leder P.;
RT   "The mouse formin (Fmn) gene: genomic structure, novel exons, and
RT   genetic mapping.";
RL   Genomics 39:303-311(1997).
RN   [3]
RP   PHOSPHORYLATION.
```

```
RX   MEDLINE=93296176; PubMed=8516300;
RA   Vogt T.F., Jackson-Grusby L., Rush J., Leder P.;
RT   "Formins: phosphoprotein isoforms encoded by the mouse limb deformity
RT   locus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:5554-5558(1993).
CC   -!- FUNCTION: Is important for the morphogenesis of limb and kidney
CC       and may have a function in differentiated cells or may be involved
CC       in maintaining specific differentiated states.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=5;
CC         Comment=Additional isoforms seem to exist;
CC       Name=IA;
CC         IsoId=Q05860-1; Sequence=Displayed;
CC       Name=IB;
CC         IsoId=Q05860-2; Sequence=VSP_001569;
CC       Name=II;
CC         IsoId=Q05860-3; Sequence=VSP_001570;
CC       Name=III;
CC         IsoId=Q05860-4; Sequence=VSP_001571, VSP_001572;
CC       Name=IV;
CC         IsoId=Q05859-1; Sequence=External;
CC   -!- TISSUE SPECIFICITY: It is present in the adult kidney, testis,
CC       limb, ovary, brain, small intestine, salivary gland and harderian
CC       gland. It is present throughout the embryo.
CC   -!- DEVELOPMENTAL STAGE: In the developing limb bud, the protein is
CC       expressed in the apical ectodermal ridge and the mesenchymal
CC       compartment, predominantly in the posterior region. During kidney
CC       morphogenesis, expression is initially restricted to the
CC       epithelial compartment of the pronephros and mesonephros.
CC   -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC   -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC   -!- SIMILARITY: Belongs to the formin homology family. Cappuccino
CC       subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X53599; CAA37668.1; -.
DR   PIR; S11515; S11515.
DR   MGD; MGI:101815; Fmn.
DR   InterPro; IPR003104; FH2.
DR   InterPro; IPR001265; Formin.
DR   Pfam; PF02181; FH2; 1.
DR   PRINTS; PR00828; FORMIN.
DR   SMART; SM00498; FH2; 1.
KW   Nuclear protein; Developmental protein; Alternative splicing;
KW   Phosphorylation; Coiled coil.
FT   DOMAIN      723    792       COILED COIL (POTENTIAL).
FT   DOMAIN      870    970       FH1 (PRO-RICH).
FT   DOMAIN      985   1426       FH2.
```

```
FT   DOMAIN      1305   1378       COILED COIL (POTENTIAL).
FT   DOMAIN       198    203       POLY-SER.
FT   DOMAIN       861    864       POLY-SER.
FT   DOMAIN       885    892       POLY-PRO.
FT   DOMAIN       911    925       POLY-PRO.
FT   DOMAIN       929    940       POLY-PRO.
FT   DOMAIN       951    962       POLY-PRO.
FT   DOMAIN       966    970       POLY-PRO.
FT   DOMAIN       977    981       POLY-SER.
FT   VARSPLIC    1252   1287       Missing (in isoform IB).
FT                                 /FTId=VSP_001569.
FT   VARSPLIC     625    722       Missing (in isoform II).
FT                                 /FTId=VSP_001570.
FT   VARSPLIC     626    627       IA -> SV (in isoform III).
FT                                 /FTId=VSP_001571.
FT   VARSPLIC     628   1468       Missing (in isoform III).
FT                                 /FTId=VSP_001572.
SQ   SEQUENCE   1468 AA;  163809 MW;  EF2FB1E9CA9DAF43 CRC64;

  Query Match             14.4%;  Score 266;  DB 1;  Length 1468;
  Best Local Similarity   17.3%;  Pred. No. 1.7;
  Matches  102;  Conservative   40;  Mismatches   97;  Indels  350;  Gaps   24;

Qy          5 DQSPLVSSSDSPPR-----PQPAFKYQFVREPEDEEE-----------------EEEEE 41
              ||||      :| |     | |  |        | ||                   : | :
Db        685 DQSPTEQDDRTPGRLQAVWPPPKTK--------DTEEKVGLKYTEAEYQAAILHLKREHK 736

Qy         42 EEDE----------------------DEDLEEL-EVLERK----------------- 58
              || |                        :| :| | : ||::
Db        737 EEIETLQAQFELKTFHIRGEHALVTARLEEAIENLKQQLEKRREGCEEMRDVCISTDDDC 796

Qy         59 -PAA------------------------------------------------------- 61
               | |
Db        797 SPKAFRNVCIQTDRETFLKPCDAESKATRSSQIVPKKLTISLTQLSPSKDSKDIHAPFQT 856

Qy         62 --GLSA---------APVPTAPAAGAPLMDFGNDFVPPAP-----RGPLPAAPPVAPERQ 105
                | |:         || || |    ||       :|| |      |||| |||: |
Db        857 REGTSSSSQQKISPPAP-PTPPPLPPPL-------IPPPPPLPPGLGPLPPAPPIPP--- 905

Qy        106 PSWDPSPVSSTVPAPSPLSAAAVSPSKLPEDDEPPARPPPPPPA--------SVSPQAEP 157
                   |||   | | |       |: :|  | ||  ||||||          |  |   |
Db        906 ----VCPVSPPPPPPPP------PPTPVPPSDGPPPPPPPPPPLPNVLALPNSGGPPPPP 955

Qy        158 VWTPPAP--APAAPPSTSV--------------------VDLLYWRDIKKTGVVFGASLF 195
                 || |  ||  ||  |                     :  |||
Db        956 PPPPPPPGLAPPPPPGLSFGLSSSSSQYPRKPAIEPSCPMKPLYWTR------------- 1002

Qy        196 LLLSLTVFSIVSVTAYIALALLSVTISFRIYKGVIQAIQKSDEGHP----------FRAY 245
                                              ||   || :  |           |
Db       1003 ---------------------------------IQINDKSQDAAPTLWDSLEEPHIRDT 1028

Qy        246 LESEVAISEELVQ-----------------------KYSNSA---LGHVNCTIKELRR 277
               | |   |::  |                       | | :    :  ::  :|::::
Db       1029 SEFEYLFSKDTTQQKKKPLSEAYEKKNKVKKIIKLLDGKRSQTVGILISSLHLEMKDIQQ 1088

Qy        278 -LFLVDDLVDSLKFAVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG 336
               :| |||                            ::: | ::  :||
```

```
Db       1089 AIFTVDD----------------------------SVVDLETLAALYE---------- 1108

Qy        337 LANKNVKDAMAKIQ------------------------AKIPGLKRKAE 361
                |:  :| : ||:                          |:||    :|:
Db       1109 --NRAQEDELTKIRKYYETSKEEDLKLLDKPEQFLHELAQIPNFAERAQ 1155
```

RESULT 14
NCR2_HUMAN
ID   NCR2_HUMAN   STANDARD;   PRT; 2517 AA.
AC   Q9Y618; O00613; O15416; Q13354; Q9Y5U0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Nuclear receptor co-repressor 2 (N-CoR2) (Silencing mediator of
DE   retinoic acid and thyroid hormone receptor) (SMRT) (SMRTe) (Thyroid-,
DE   retinoic-acid-receptor-associated co-repressor) (T3 receptor-
DE   associating factor) (TRAC) (CTG repeat protein 26).
GN   NCOR2 OR CTG26.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM SMRT).
RC   TISSUE=Pituitary;
RX   MEDLINE=99178941; PubMed=10077563;
RA   Ordentlich P., Downes M., Xie W., Genin A., Spinner N.B., Evans R.M.;
RT   "Unique forms of human and mouse nuclear receptor corepressor SMRT.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:2639-2644(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM SMRT).
RC   TISSUE=Cervical adenocarcinoma;
RX   MEDLINE=99199215; PubMed=10097068;
RA   Park E.J., Schroen D.J., Yang M., Li H., Li L., Chen J.D.;
RT   "SMRTe, a silencing mediator for retinoid and thyroid hormone
RT   receptors-extended isoform that is more related to the nuclear
RT   receptor corepressor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:3519-3524(1999).
RN   [3]
RP   SEQUENCE OF 1023-2517 FROM N.A.
RC   TISSUE=Cervical adenocarcinoma;
RX   MEDLINE=96008552; PubMed=7566127;
RA   Chen J.D., Evans R.M.;
RT   "A transcriptional co-repressor that interacts with nuclear hormone
RT   receptors.";
RL   Nature 377:454-457(1995).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM TRAC-1).
RC   TISSUE=Fetal liver;
RX   MEDLINE=96408715; PubMed=8813722;
RA   Sande S., Privalsky M.L.;
RT   "Identification of TRACs (T3 receptor-associating cofactors), a family
RT   of cofactors that associate with, and modulate the activity of,
RT   nuclear hormone receptors.";
RL   Mol. Endocrinol. 10:813-825(1996).
RN   [5]

```
RP   SEQUENCE OF 428-613 FROM N.A.
RC   TISSUE=Brain cortex;
RX   MEDLINE=97369492; PubMed=9225980;
RA   Margolis R.L., Abraham M.R., Gatchell S.B., Li S.-H., Kidwai A.S.,
RA   Breschel T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A.;
RT   "cDNAs with long CAG trinucleotide repeats from human brain.";
RL   Hum. Genet. 100:114-122(1997).
RN   [6]
RP   INTERACTION WITH MINT.
RX   MEDLINE=21231190; PubMed=11331609;
RA   Shi Y., Downes M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
RA   Hon M., Evans R.M.;
RT   "Sharp, an inducible cofactor that integrates nuclear receptor
RT   repression and activation.";
RL   Genes Dev. 15:1140-1151(2001).
RN   [7]
RP   INTERACTION WITH HDAC10.
RX   MEDLINE=21839031; PubMed=11739383;
RA   Fischer D.D., Cai R., Bhatia U., Asselbergs F.A.M., Song C., Terry R.,
RA   Trogani N., Widmer R., Atadja P., Cohen D.;
RT   "Isolation and characterization of a novel class II histone
RT   deacetylase, HDAC10.";
RL   J. Biol. Chem. 277:6656-6666(2002).
CC   -!- FUNCTION: Mediates the transcriptional repression activity of some
CC       nuclear receptors by promoting chromatin condensation, thus
CC       preventing access of the basal transcription.
CC   -!- SUBUNIT: Interacts with HDAC7 (By similarity). Forms a large
CC       corepressor complex that contains SIN3A/B and histone deacetylases
CC       HDAC1 and HDAC2. This complex associates with the thyroid (TR) and
CC       the retinoid acid receptors (RAR) in the absence of ligand, and
CC       may stabilize their interaction with TFIIB. The SRMT isoform
CC       interacts with HDAC10. Interacts with MINT.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=SMRT; Synonyms=TRAC-2;
CC         IsoId=Q9Y618-1; Sequence=Displayed;
CC       Name=TRAC-1;
CC         IsoId=Q9Y618-2; Sequence=VSP_003412, VSP_003413;
CC         Note=Contains only the C-terminal receptor-interacting domain
CC         and acts as an antirepressor;
CC   -!- TISSUE SPECIFICITY: Ubiquitous. High levels of expression are
CC       detected in lung, spleen and brain.
CC   -!- INDUCTION: Regulated during cell cycle progression.
CC   -!- DOMAIN: The N-terminal region contains repression functions that
CC       are divided into three independent repression domains (RD1, RD2
CC       and RD3). The C-terminal region contains the nuclear receptor-
CC       interacting domains that are divided in two separate interaction
CC       domains (ID1 and ID2).
CC   -!- DOMAIN: The two interaction domains (ID) contain a conserved
CC       sequence referred to as the CORNR box. This motif is required and
CC       sufficient to permit binding to unliganded TR and RARS. Sequences
CC       flanking the CORNR box determine nuclear hormone receptor
CC       specificity.
CC   -!- SIMILARITY: Contains 1 SANT-A domain.
CC   -!- SIMILARITY: Contains 1 Myb-like domain.
CC   -!- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
```

```
CC       FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF113003; AAD20946.1; -.
DR   EMBL; AF125672; AAD22973.1; -.
DR   EMBL; U37146; AAC50236.1; -.
DR   EMBL; S83390; AAB50847.1; -.
DR   EMBL; U80750; AAB91446.1; -.
DR   PIR; S60255; S60255.
DR   PDB; 1KKQ; 27-FEB-02.
DR   TRANSFAC; T04689; -.
DR   Genew; HGNC:7673; NCOR2.
DR   MIM; 600848; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR   InterPro; IPR001005; Myb_DNA_binding.
DR   Pfam; PF00249; myb_DNA-binding; 2.
DR   SMART; SM00717; SANT; 2.
DR   PROSITE; PS50090; MYB_3; 1.
KW   Nuclear protein; Transcription regulation; DNA-binding; Repressor;
KW   Coiled coil; Alternative splicing; 3D-structure.
FT   DOMAIN      174    215       COILED COIL (POTENTIAL).
FT   DOMAIN      254    312       INTERACTION WITH SIN3A/B (BY SIMILARITY).
FT   DNA_BIND    429    474       SANT-A (POTENTIAL).
FT   DNA_BIND    613    657       MYB.
FT   DOMAIN      522    561       COILED COIL (POTENTIAL).
FT   DOMAIN      778    820       PRO-RICH.
FT   DOMAIN     2139   2143       CORNR BOX OF ID1.
FT   DOMAIN     2342   2346       CORNR BOX OF ID2.
FT   DOMAIN      494    510       POLY-GLN.
FT   DOMAIN      682    685       POLY-LYS.
FT   DOMAIN      994   1002       POLY-PRO.
FT   DOMAIN     1384   1389       POLY-PRO.
FT   DOMAIN     1842   1846       POLY-GLY.
FT   DOMAIN     2479   2482       POLY-PRO.
FT   VARSPLIC      1   1702       Missing (in isoform TRAC-1).
FT                                /FTId=VSP_003412.
FT   VARSPLIC   2353   2398       Missing (in isoform TRAC-1).
FT                                /FTId=VSP_003413.
FT   CONFLICT      7      7       L -> P (IN REF. 2).
FT   CONFLICT    295    295       K -> E (IN REF. 2).
FT   CONFLICT    309    309       L -> W (IN REF. 2).
FT   CONFLICT    352    352       MISSING (IN REF. 2).
FT   CONFLICT    365    365       A -> P (IN REF. 2).
FT   CONFLICT    612    613       SS -> EF (IN REF. 5).
FT   CONFLICT    711    711       S -> T (IN REF. 2).
FT   CONFLICT    724    740       MISSING (IN REF. 2).
FT   CONFLICT    787    796       RRTSRAPIEP -> PEDIPAPTES (IN REF. 2).
FT   CONFLICT    804    804       G -> L (IN REF. 2).
FT   CONFLICT    814    814       S -> F (IN REF. 2).
```

```
FT   CONFLICT    817    817       A -> S (IN REF. 2).
FT   CONFLICT    889    889       G -> R (IN REF. 2).
FT   CONFLICT   1023   1030       SRSPAPPA -> MEAWDAHP (IN REF. 3).
FT   CONFLICT   1034   1034       A -> AEKPVFFPA (IN REF. 2).
FT   CONFLICT   1894   1894       K -> T (IN REF. 4).
FT   CONFLICT   2494   2494       P -> A (IN REF. 4).
SQ   SEQUENCE   2517 AA;  274031 MW;  F5805C01761258C0 CRC64;

  Query Match             14.0%;  Score 259.8;  DB 1;  Length 2517;
  Best Local Similarity   11.9%;  Pred. No. 10;
  Matches  127;  Conservative   43;  Mismatches  102;  Indels  792;  Gaps   29;

Qy          6 QSPLVSSSDSPPRPQPAFKYQFVREPEDEEEEEEEEEDEDEDL---------------- 49
              | |:       ||     | :  :| | |:|||: | |:: |||
Db        509 QQPM-------PRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEK 561

Qy         50 ------------------------------------------------------EELE 53
                                                                    ||:|
Db        562 EAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEME 621

Qy         54 VLER------------------------------------------------------- 57
                ::
Db        622 TAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARR 681

Qy         58 ------------------------------------------------------------ 57

Db        682 KKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPAT 741

Qy         58 --------------------------KPAAGLSA-------------------APVPTAP 72
                                        || | | |                   | | :
Db        742 VNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASE 801

Qy         73 AAGAPLMDFGNDFVPPAPRGPLPAAPPVAP-----------------ERQPSWDPS--- 111
              | |||         ||||  |  | ||| |                  |::|   |:
Db        802 ATGAPT-------PPPAPPSP-SAPPPVVPKEEKEEETAAAPPVEEGEEQKP---PAAEE 850

Qy        112 ----------PVSSTV------------------------------------------- 117
                       || |
Db        851 LAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSG 910

Qy        118 ----------------------------PAPSPLS---------------------A 125
                                          | || |:                      |
Db        911 APQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRA 970

Qy        126 AAVSPSKLPEDDEP---------PARPPPPPPASVSPQAE------------------- 156
              ||: | :: :  ||         || | |||| :: |:::
Db        971 AAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPA 1030

Qy        157 ----------------PVWT----------------PPAPAPAA---------------- 168
                             | ||                | || |:|
Db       1031 DKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHD 1090

Qy        169 --------PPSTS-------------------------------------------VV 175
                    ||: |                                           :
Db       1091 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 1150

Qy        176 DLLYWRDIKKTGVVFGASLFLLLSLTVFSIV----------------------SVTA 210
```

```
Db       1151 GLPLPMDPKK--------------LAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196

Qy        211 YIALALL---SVT-------------ISFR-----------IYKGVIQAI---------- 233
Db       1197 GTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLD 1256

Qy        234 ----------------------------------QKSD---------------------EG 239
Db       1257 RGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEG 1316

Qy        240 HPFRA-------------------------------------------------YLESEVA 251
Db       1317 RVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAK 1376

Qy        252 I------------SEELVQKYSNSALGHVN---------CTIKELRRLFLVDDLVDSLKF 290
Db       1377 LLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGR------------- 1423

Qy        291 AVLMWVFTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLGLANKNVK------- 343
Db       1424 -------------------------SIHEIPREELRHTPE----LPLAPRPLKEGSITQG 1454

Qy        344 -----------------------------------DAMAKIQA 351
Db       1455 TPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARA 1498
```

```
RESULT 15
DIA1_HUMAN
ID   DIA1_HUMAN     STANDARD;      PRT;  1248 AA.
AC   O60610; Q9UC76;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Diaphanous protein homolog 1 (Diaphanous-related formin 1) (DRF1).
GN   DIAPH1 OR DIAP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98028756; PubMed=9360932;
RA   Lynch E.D., Lee M.K., Morrow J.E., Welcsh P.L., Leon P.E., King M.-C.;
RT   "Nonsyndromic deafness DFNA1 associated with mutation of a human
RT   homolog of the Drosophila gene diaphanous.";
RL   Science 278:1315-1318(1997).
RN   [2]
RP   SEQUENCE OF 218-817 FROM N.A.
RC   TISSUE=Ovarian carcinoma;
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
```

```
RA   Ninomiya K., Iwayanagi T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 727-765 AND 1121-1145.
RC   TISSUE=Platelet;
RX   MEDLINE=95255215; PubMed=7737110;
RA   Reinhard M., Giehl K., Abel K., Haffner C., Jarchau T., Hoppe V.,
RA   Jockusch B.M., Walter U.;
RT   "The proline-rich focal adhesion and microfilament protein VASP is a
RT   ligand for profilins.";
RL   EMBO J. 14:1583-1589(1995).
CC   -!- FUNCTION: Binds to GTP-bound form of Rho and to profilin. Acts in
CC       a Rho-dependent manner to recruit profilin to the membrane, where
CC       it promotes actin polymerization. It is required for cytokinesis,
CC       stress fiber formation, and transcriptional activation of the
CC       serum response factor. DFR proteins couple Rho and Src tyrosine
CC       kinase during signaling and the regulation of actin dynamics (By
CC       similarity). In hearing it may play a role in the regulation of
CC       actin polymerization in hair cells.
CC   -!- TISSUE SPECIFICITY: Expressed in brain, heart, placenta, lung,
CC       kidney, pancreas, liver, skeletal muscle and cochlea.
CC   -!- DOMAIN: DRFs are regulated by intramolecular GBD-DAD binding where
CC       Rho-GTP activates the DRFs by disrupting the GBD-DAD interaction
CC       (By similarity).
CC   -!- DISEASE: Defects in DIAPH1 are a cause of autosomal dominant
CC       nonsyndromic sensorineural deafness 1 (DFNA1) [MIM:124900].
CC   -!- SIMILARITY: Contains 1 GTPase-binding (GBD) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 1 (FH1) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 2 (FH2) domain.
CC   -!- SIMILARITY: Contains 1 Formin homology 3 (FH3) domain.
CC   -!- SIMILARITY: Contains 1 DRF autoregulatory (DAD) domain.
CC   -!- SIMILARITY: Belongs to the formin homology family. Diaphanous
CC       subfamily.
CC   -!- CAUTION: Ref.2 sequence differs from that shown in that it seems
CC       to include intronic sequence.
CC   -!- DATABASE: NAME=Hereditary hearing loss homepage;
CC       NOTE=Gene page;
CC       WWW="http://www.uia.ac.be/dnalab/hhh/".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF051782; AAC05373.1; -.
DR   EMBL; AK023345; BAB14533.1; ALT_SEQ.
DR   Genew; HGNC:2876; DIAPH1.
DR   MIM; 602121; -.
DR   MIM; 124900; -.
DR   GO; GO:0007605; P:hearing; TAS.
DR   InterPro; IPR003104; FH2.
DR   Pfam; PF02181; FH2; 1.
DR   SMART; SM00498; FH2; 1.
```

```
KW   Coiled coil; Repeat; Deafness.
FT   DOMAIN       63    260       GBD.
FT   DOMAIN      157    457       FH3.
FT   DOMAIN      460    563       COILED COIL (POTENTIAL).
FT   DOMAIN      588    743       FH1 (PRO-RICH).
FT   DOMAIN      748   1190       FH2.
FT   DOMAIN     1015   1172       COILED COIL (POTENTIAL).
FT   DOMAIN     1173   1187       DAD.
FT   DOMAIN     1189   1192       ARG/LYS-RICH (BASIC).
FT   CONFLICT    804    804       T -> TSKA (IN REF. 2).
FT   CONFLICT   1132   1133       RK -> AE (IN REF. 3).
SQ   SEQUENCE   1248 AA;  138978 MW;  ED1F5147CFF9A886 CRC64;

  Query Match             14.0%;  Score 259.7;  DB 1;  Length 1248;
  Best Local Similarity   15.6%;  Pred. No. 1.9;
  Matches  143;  Conservative   49;  Mismatches  114;  Indels  613;  Gaps   40;

Qy          1 MEDLD----------------------QSPLVSSSDSPPRPQ----------------- 20
              |:| :                      |  |:  :|   |||
Db        378 MDDFNEVFQILLNTVKDSKAEPHFLSILQHLLLVRNDYEARPQYYKLIEECISQIVLHKN 437

Qy         21 ---PAF------------------------------------------------------ 23
                | |
Db        438 GADPDFKCRHLQIEIEGLIDQMIDKTKVEKSEAKAAELEKKLDSELTARHELQVEMKKME 497

Qy         24 -----KYQFVREPEDEEEEEEEEEDEDEDLE---------------ELEVLERKPAAGL 63
                 | | ::  :|    |:::   | :|||               |||   :|  | |
Db        498 SDFEQKLQDLQGEKDALHSEKQQIATEKQDLEAEVSQLTGEVAKLTKELED-AKKEMASL 556

Qy         64 SA------------APVPTA-----------PAAGAPLMDFGNDFVPPAPRGPLPAAPP- 99
              ||            |||| |           |   ||    |:   || |  | |  ||
Db        557 SAAAITVPPSVPSRAPVPPAPPLPGDSGTIIPPPPAP----GDSTTPPPPPPPPPPPPL 612

Qy        100 -----VAPERQPSWDPSPVS--STVPAPSPL--SAAAVSPSKLPE----DDEPP----AR 142
                  ::|       | |:|  :|:| | ||       ||| ||        ||    ||
Db        613 PGGTAISP-------PPPLSGDATIPPPPPLPEGVGIPSPSSLPGGTAIPPPPPLPGSAR 665

Qy        143 -PPPPPP----ASVSPQAEPV-----------------WTPPAPAPAAP---------- 169
               ||||||    | : |   |:                  || | |  |
Db        666 IPPPPPPLPGSAGIPPPPPPLPGEAGMPPPPPPLPGGPGIPPPPPFPGGPGIPPPPPGMG 725

Qy        170 -----------PSTSVV--------------------------DL---LYWRDIK---- 184
                        |:  |:                          ||    :|  :|
Db        726 MPPPPPFGFGVPAAPVLPFGLTPKKLYKPEVQLRRPNWSKLVAEDLSQDCFWTKVKEDRF 785

Qy        185 ---------------------------------------------KT------------ 186
                                                          ||
Db        786 ENNELFAKLTLTFSAQTKTKKDQEGGEEKKSVQKKKVKELKVLDSKTAQNLSIFLGSFRM 845

Qy        187 ---------------------------------------------------GVVFG- 191
                                                                |||  |
Db        846 PYQEIKNVILEVNEAVLTESMIQNLIKQMPEPEQLKMLSELKDEYDDLAESEQFGVVMGT 905

Qy        192 ---------ASLFLLLSLTVFS---------IVSVTAYIALA-----------LLSVTI 221
                     | || |     ||         ||||||    |            || :|:
Db        906 VPRLRPRLNAILFKL----QFSEQVENIKPEIVSVTA----ACEELRKSESFSNLLEITL 957
```

```
Qy        222 ----------------SFRI----------------------------YKGVIQAIQKS 236
                              | |                               |  |::
Db        958 LVGNYMNAGSRNAGAFGFNISFLCKLRDTKSTDQKMTLLHFLAELCENDYPDVLKF---P 1014

Qy        237 DEGHPFRAYLESEVAISEELVQK----------------------------------- 259
              ||     |::|    :| | :||
Db       1015 DE----LAHVEKASRVSAENLQKNLDQMKKQISDVERDVQNFPAATDEKDKFVEKMTSFV 1070

Qy        260 ------YSNSALGHVN--CTIKEL----------------------RRLFL-------- 280
                   |:   : | |     |||                      | :||
Db       1071 KDAQEQYNKLRMMHSNMETLYKELGEYFLFDPKKLSVEEFFMDLHNFRNMFLQAVKENQK 1130

Qy        281 ------------------------------------------VDDLVDSLKFAVLMWV 296
                                                       :| |:::|:
Db       1131 RRKTEEKMRRAKLAKEKAEKERLEKQQKREQLIDMNAEGDETGVMDSLLEALQ------- 1183

Qy        297 FTYVGALFNGLTLLILALISLFSVPVIYERHQAQIDHYLG------LANKNVK-DAMAKI 349
                 || |                   :|   | :   |      ||::  | |||| :
Db       1184 ---SGAAFR----------------RKRGPRQANRKAGCAVTSLLASELTKDDAMAAV 1222

Qy        350 QAKI-------PGLKRKAE 361
               ||:       | :  :|:
Db       1223 PAKVSKNSETFPTILEEAK 1241
```

Search completed: September 29, 2004, 18:52:12
Job time : 20.0208 secs